Department of Civil Engineering

University of Strathclyde

# Environmental Forensic Investigation of Coal Tars from Former Manufactured Gas Plants

A Thesis presented for the Degree of Doctor of Philosophy

by

Laura McGregor

2012

This thesis is the result of the author's original research. It has been composed by the author and has not been previously submitted for examination which has led to the award of a degree.

The copyright of this thesis belongs to the author under the terms of the United Kingdom Copyright Acts as qualified by University of Strathclyde Regulation 3.50. Due acknowledgement must always be made of the use of any material contained in, or derived from, this thesis.

Signed:

Date:

# Abstract

Former manufactured gas plants (FMGPs) are a ubiquitous source of environmental contamination. The process of gas production created a number of by-products, including coal tar and ammoniacal liquor. Coal tar contains a complex mixture of organic and inorganic compounds, many of which are toxic and carcinogenic. It is estimated that over $3000\pm1000$ FMGPs exist in the United Kingdom alone, yet there are few recent publications detailing the analysis of coal tars. The complex composition of coal tar is known to vary due to a number of factors, including production method, temperature and coal type, making the analysis and interpretation of such samples extremely challenging.

Environmental forensics is concerned with the source, fate and transport of contaminants. The introduction of recent legislation such as the EU Environmental Liabilities Directive 2004/35/EC, which promotes the "polluters pay" policy, has encouraged the development of accurate and robust scientific methods for the identification of contaminants. Analytical instrumentation is constantly evolving, thus new protocols to trace the origin of contamination must also be developed to utilise these technological advances. Two-dimensional gas chromatography (GCxGC) and compound specific isotope analysis (CSIA) are two examples of advanced analytical instruments which have the potential to aid source identification. GCxGC provides enhanced separation of complex mixtures compared to conventional gas chromatographic techniques, while CSIA allows chemically identical contaminants to be compared based on their isotopic composition.

In this study, preliminary research investigated the isotopic composition of coal tars, as this

technique is currently a major tool for source apportionment in environmental forensics. However, the results demonstrated that the similar nature of British coals used for gas production at the investigated sites produced similar isotopic values in the resultant tars, making the technique redundant in this case. Therefore, the potential of GCxGC was investigated for ultra resolution chemical fingerprinting of coal tars.

Traditionally, chemical fingerprinting of complex mixtures, such as coal tar, is performed using a tiered approach including rigorous sample preparation steps and analysis by multiple instruments. In this work, a new, single-step analytical procedure was developed for the analysis of coal tars by GCxGC. Automated sample extraction techniques combined with GCxGC analyses were employed to provide detailed chemical fingerprinting in a fast, yet accurate, manner. This research represents a major advance in knowledge of compositional variation within coal tars. The enhanced separation of GCxGC provides vast quantities of chemical data which can be difficult to interpret without statistical methods. A multivariate statistical model was developed to provide process-specific classification of coal tars. The statistical model was validated through use of a blind study, indicating that process-specific apportionment of coal tars was achievable. Based on these results, the application of the single-step procedure for environmental forensics on a commercial basis was evaluated.

# Acknowledgements

First and foremost, I would like to thank my supervisors, Prof. Bob Kalin and Prof. Niamh Nic Daéid, for all their support, motivation and guidance throughout my PhD research. Special thanks go to Dr Caroline Gauchotte (my unofficial supervisor) for being there with advice and generally being a good friend. If it wasn't for her encouragement to stop running experiments and start writing, I'd probably still be working away in the lab. Another thank you goes to Russell Thomas at Parsons Brinckerhoff for suppling coal tar samples for this research and for his constant enthusiasm for the subject. Thank you also to National Grid (especially Paddy Daly) for authorising analysis of the coal tar samples. I would also like to thank other members of the environmental research group, both past and present, for making my PhD an enjoyable experience, especially my office-mate Lisa, Antoine, Emma, Gráinne, Seánín and Steph, as well as other members of academic staff and the ladies in the main office. I would also like to express my gratitude to Victoria Pitstra (2011 MSc graduate at the University of Strathclyde) for investigating the cost implications of the ultra resolution chemical fingerprinting method as described in this thesis. A big thank you goes to my family for their emotional and financial support throughout my university studies. Thank you to both my family and friends (Becca, Emily, Donna, Kayley, Duncan, Douglas, Laura and the Sarahs) who listened to my, seemingly endless, rants about broken lab equipment, they all now know far more than they ever wanted to about a certain two-dimensional gas chromatograph (aka Sneezy). I'm sure there are many others I could go on to thank, but I've been told that's a cheap way to pad out a thesis!

# Posters, Presentations and Publications

**Peer-reviewed Publications**

L. A. McGregor, C. Gauchotte-Lindsay, N. Nic Daéid, R. Thomas, P. Daly, R. M. Kalin, Ultra Resolution Chemical Fingerprinting of Dense Non-Aqueous Phase Liquids from Manufactured Gas Plants by Reversed Phase Comprehensive Two-Dimensional Gas Chromatography, *J. Chromatogr. A*, 2011, 1218, 4755-4763.

L. A. McGregor, C. Gauchotte-Lindsay, N. Nic Daéid, R. Thomas, R. M. Kalin, Multivariate statistical methods for the environmental forensic classification of coal tars from former manufactured gas plants, *Env. Sci. Technol.,* Just Accepted, February 2012.

**Presentations**

L. A. McGregor, C. Gauchotte, P. Richards, S. Kerr, H. Gray, N. N. Daéid, R. Thomas, P. Daly and R. M. Kalin. 2011. Forensic Analysis of Coal Tar from Former Manufactured Gas Plants, *Sustainability Live, Brownfield Expo*, 24th-26th May 2011; Birmingham, U.K.

L. A. McGregor, C. Gauchotte, N. N. Daéid, R. Thomas, P. Daly and R. M. Kalin. 2011. Next Generation Chemical Fingerprinting: Case Study of Coal Tar from Former Manufactured Gas Plants. 2011. *International Network of Environmental Forensics (INEF) conference*, 25th-27th July 2011; Cambridge, U.K. ***1st place in the student presentation competition.***

**Posters**

L. A. McGregor, C. Gauchotte, A. Assal, R. M. Kalin and N. Nic Daéid. 2009. Development of a Quality Control Procedure for Compound Specific Hydrogen Isotope Ratio Mass Spectrometry, *International Network of Environmental Forensics (INEF) conference,* 31st August-2nd September 2009; Calgary, Canada.

L. N. Kates, L. A. McGregor, C. Gauchotte, A. Assal, R. M. Kalin and N. Nic Daéid. 2009. Comprehensive Two-Dimensional Gas Chromatography Time of Flight Mass Spectrometry: Investigating Degradation Trends of Dense Non-Aqueous Phase Liquids, *International Network of Environmental Forensics (INEF) conference,* 31st August-2nd September 2009; Calgary, Canada.

L. A. McGregor, L. N. Kates, C. Gauchotte, A. Assal and R. M. Kalin. 2009. Comprehensive Two-Dimensional Gas Chromatography Time of Flight Mass Spectrometry for Environmental Forensics Applications: Fingerprinting of Dense Non-aqueous Phase Liquids, *European Academy of Forensic Science (EAFS) conference*, 9th − 11th September 2009; University of Strathclyde, Glasgow.

L. A. McGregor, L. N. Kates, C. Gauchotte, N. N. Daéid and R. M. Kalin. 2010. Comprehensive Two-Dimensional Gas Chromatography Time of Flight Mass Spectrometry for Environmental Forensics Applications: Case Study of a Former Manufactured Gas Plant, *Glasgow Research Partnership in Engineering (GRPE) review event*, 24th June 2010; University of Strathclyde, Glasgow.

# Contents

# List of Figures

# List of Tables

# Abbreviations

| | |
|---|---|
| ACE | acenaphthene |
| ACY | acenaphthylene |
| ANT | anthracene |
| ASE | accelerated solvent extraction |
| BaA | benz[a]anthracene |
| BaP | benzo[a]pyrene |
| BbF | benzo[b]fluoranthene |
| BkF | benzo[k]fluoranthene |
| BP | benzo[g,h,i]perylene |
| cf | cubic feet |
| CHR | chrysene |
| CI | confidence interval |
| CO | coke oven |
| CR | creosote oil |
| CSIA | compound specific isotope analysis |
| CWG | carburetted water gas |
| CVR | continuous vertical retort |
| DBA | dibenzo[a,h]anthracene |

| | |
|---|---|
| DNAPL | dense non-aqueous phase liquid |
| FLU | fluorene |
| FLT | fluoranthene |
| FMGP | former manufactured gas plant |
| GC | gas chromatography |
| GC-FID | gas chromatography – flame ionisation detector |
| GC-MS | gas chromatography – mass spectrometry |
| GC-C-IRMS | gas chromatography - combustion isotope ratio mass spectrometry |
| GCxGC | two-dimensional gas chromatography |
| id | internal diameter |
| IP | indeno[1,2,3-cd]pyrene |
| IRMS | isotope ratio mass spectrometer |
| HCA | hierarchical cluster analysis |
| HR | horizontal retort |
| mg/kg | milligrams per kilogram |
| MGP | manufactured gas plant |
| N | naphthalene |
| PAH | polycyclic aromatic hydrocarbon |
| PCA | principal component analysis |
| PHE | phenanthrene |
| RSD | relative standard deviation |
| TOFMS | time of flight mass spectrometry |
| US EPA | United States Environmental Protection Agency |
| VR | vertical retort |

# Chapter 1

# Introduction

## 1.1 Background

Environmental forensics is concerned with the scientific and incontrovertible determination of the responsible party to an incidence of pollution. Robust analytical methods are required to remove any obstacles regarding implementation of recent legislation, such as the Stockholm Convention and the EU Environmental Liabilities Directive 2004/35/EC. However, accurate characterisation of a contaminated site is not required solely to aid legal proceedings; it is also a crucial step in assessing the toxicity of a site and thus the employment of appropriate remediation techniques.

The complex nature of environmental samples allows a chemical fingerprint to be generated, which can then be used to compare contaminants with potential sources. Due to technological developments in current instrumentation, an updated form of chemical fingerprinting would advance the science of environmental forensic analyses of complex mixtures, such as petroleum products or coal tar. Advanced analytical techniques such as comprehensive two-dimensional gas chromatography (GCxGC) and gas chromatography coupled with isotope ratio mass spectrometry (GC-C-IRMS) have proven to be successful in a variety of fields, but their full capabilities have yet to be utilised in an accurate and precise method of chemical fingerprinting. The use of such advanced instrumentation could potentially overcome the limitations of traditional techniques. For example, the resolving power of GCxGC is an order of magnitude greater than that of traditional gas chromatographic methods allowing enhanced separation of complex mixtures such as crude oils. Moreover,

GC-C-IRMS is capable of distinguishing between chemically identical compounds based on isotopic composition, thus can provide an additional level of information for source apportionment.

The potential of these techniques will be measured through analysis of coal tars obtained from former manufactured gas plants (FMGPs) in the United Kingdom. FMGP sites are ubiquitous in the United Kingdom, thus are a major source of contamination due to the complex by-products which were released into the environment throughout the years of operation. Due to the vast range of compounds found in such samples they are an ideal choice for testing the capabilities of the new method of chemical fingerprinting.

## 1.2   Research Aim and Objectives

The aim of this research is to investigate an expanded form of chemical fingerprinting using advanced analytical techniques to allow more accurate and precise source apportionment and to gain knowledge on the chemical composition of coal tars. To achieve the aims of this project, the following research objectives were identified:

- To evaluate the potential of compound specific isotope analysis (CSIA) as a tool for source differentiation of coal tars.

- To optimise the use of two-dimensional gas chromatography (GCxGC) for the analysis of coal tars.

- To develop a fast, accurate and precise extraction procedure to allow analysis of complex coal tars by GCxGC without the need for complex fractionation processes.

- To identify the main classes of chemicals present in coal tars through tentative GCxGC identification for the purpose of chemical fingerprinting.

- To compare statistical methods of processing GCxGC datasets for source apportionment of coal tars.

## 1.3    Thesis Outline

This thesis details the development of an improved analytical procedure for source allocation of coal tars.

Chapter 2 presents an introduction to environmental forensics and the analytical methods used to trace the origin of contaminants. Chapter 2 also provides background information on the manufactured gas industry and the various forms of contamination found at the site of former gasworks. Lastly, a review of previous environmental forensic investigations at former manufactured gas plants is presented.

Chapter 3 explains the theory of operation behind the analytical and spectroscopic techniques employed in this work. The principles of gas chromatography, mass spectrometry, two-dimensional gas chromatography and isotope analysis are described in detail.

Chapter 4 presents an overview of the methodology employed in this research. The specifications and parameters of instrumental techniques are provided, as well as quality control procedures. Brief site summaries are also provided for each FMGP investigated in this study.

Chapter 5 presents the study of the carbon isotopic values of coal tars through compound specific isotope analysis. The development of a novel fractionation method using an automated extraction technique is firstly described. The carbon isotope data for the aromatic fraction of each sample is then evaluated.

Chapter 6 provides a review of relevant literature on two-dimensional gas chromatography (GCxGC) and its application in environmental forensics. A single-step approach for ultra resolution chemical fingerprinting of coal tars is presented using GCxGC. The optimisation of extraction method and GCxGC parameters are detailed.

Chapter 7 first presents the theory behind statistical data analysis and the application of these methods to environmental forensics. A variety of univariate and multivariate statistical methods were evaluated for effective interpretation of GCxGC data. An optimised statistical model for source allocation of coal tars is presented.

Chapter 8 evaluates the classification power of the statistical model developed within Chapter 7,

through source apportionment of unknown coal tar samples. The potential application of ultra resolution chemical fingerprinting for commercial environmental forensics is also discussed.

Finally, Chapter 9 summarises the major findings of this thesis and highlights the recommendations for further work.

# Chapter 2

# Literature Review

## 2.1  Introduction

Constant increases in energy demands and growing industrialisation has resulted in environmental contamination becoming a global problem, either through minor accidental leaks or chronic spills [Wang and Fingas, 1997]. Major incidents, such as the *Exxon Valdez* oil spill in the Prince William Sound, Alaska in 1989 [Boehm et al., 1998] prompted the growth of environmental forensics; an industry dedicated to the identification of contaminants in the environment. Environmental forensics is the study of the source, fate and transport of contaminants in the environment [Morrison, 2000]. This can prove highly challenging, as processes such as weathering and biological degradation will begin to alter contaminants as soon as they are released into the environment [Mudge, 2008]. Furthermore, the complex nature of most environmental contaminants, such as coal tars or petroleum, requires robust scientific methods to identify and compare samples with possible sources [O'Sullivan and Kalin, 2008].

The aim of this research was to develop an accurate and efficient analytical method for environmental forensic investigation of coal tar contamination at former manufactured gas plants. In this Chapter, the application of science in law will be discussed, as well as the impact of new environmental legislation on analytical procedures employed in environmental forensic investigations. The formation of coal tar by historic gas manufacturing processes will then be described and previous research into their complex composition will be reviewed.

## 2.2    Environmental Forensics and Legislation

The analysis of environmental samples to detect contamination is by no means a recent application; the difference for environmental forensics now being the allegation of blame, and thus remediation costs, to those found responsible for pollution [Mudge, 2008]. Furthermore, establishing the level of toxic contamination released into the environment by a given source can aid compensation claims regarding illness or damage to property relating to that contamination [Jones and Parpworth, 2004].

The introduction of environmental legislation, such as the European Environmental Liability Directive (2004/35/EC) and the United States Comprehensive Environmental Response, Compensation and Liability Act [CERCLA, 1980] which promote the "polluters pay" principle, has increased the need for advanced analytical techniques that accurately identify sources of contamination [Mudge, 2008]. For example, the European Environmental Liability Directive (2004/35/EC) states that:

"...an operator whose activity has caused the environmental damage or the imminent threat of such damage is to be held financially liable, in order to induce operators to adopt measures and develop practices to minimise the risks of environmental damage so that their exposure to financial liabilities is reduced." [DEFRA, 2004].

In other words, polluters will be forced to bear the cost of remediating any land damaged by contamination [Mudge, 2008]. The application of science in law creates a number of issues regarding the acceptability and admissibility of evidence. Forensic evidence has been used for many years within criminal law, allowing a number of legal requirements regarding the quality of scientific data to be established. For example, the Frye Rule, developed during a criminal case where a precursor to the lie detector was used to convict the defendant, was the first criterion used to moderate the quality of scientific data [Frye v. United States, D.C. circa. 1923]. The rule states that;

"...the thing from which a deduction is made must be sufficiently established to have gained general acceptance in the particular field in which it belongs" [Frye v. United States, D.C. circa. 1923].

The ambiguous nature of the phrasing in this rule caused debate within the legal field, resulting in a further, more rigid, set of rules to be developed; known as the Daubert criteria [Daubert v. Merrell Dow Pharmaceuticals, 1993].

The Daubert criteria is composed of four main questions which are used to assess the admissability of scientific evidence, namely; "(1) Does the theory or technique involve testable hypotheses? (2)

Has the theory or technique been subject to peer review and publication? (3) Are there known or potential error rates and are there standards controlling the technique's operation? (4) Is the method or technique generally accepted in the scientific community?" [Brilis et al., 2000].

Brillis et al. (2000) compared the Daubert rules to the Quality System published by the US Environmental Protection Agency (US EPA) and concluded that the quality assurance guidelines set by the US EPA cover each of the Daubert factors. However, Brillis et al. also state that given the abundant use of novel techniques in environmental forensic science, it is essential that the analytical scientist follows strict quality assurance and quality control policies and procedures. This is even more pertinent in legal cases within the United Kingdom, as UK law lacks a firm set of admissabilty rules, such as the Daubert criteria, and instead relies on a series of exclusions based on case law, with the final decision ultimately made by a judge [Fraser, 2010]. The environmental forensics industry is expanding within the UK, partly due to the landmark Corby case [Corby Group Litigation v Corby District Council, 2009]. The case involved compensation claims by four families whose children suffered illnesses, allegedly caused by contamination spread during remediation efforts at a former steel works site. The introduction of strict environmental legislation has increased the demand for environmental forensics experts to present evidence in court which can link a particular source of contamination with detrimental effects to human health or the environment [Brilis et al., 2000]. It must be noted that the guidelines for the quality of scientific data recognised by criminal law should also be applied to environmental law. In short, the protection of the environment is dependent on the quality of scientific data produced at a trial by environmental forensic experts [Brilis et al., 2000].

## 2.3   Chemical Fingerprinting

After release into the environment, contaminants are immediately subjected to weathering processes, such as dissolution or evaporation, which may alter the chemical composition [Wang and Fingas, 1997]. Consequently, accurate and reliable analytical procedures are required to allow unambiguous identification of contaminant source [Wang et al., 1999]. The complex nature of most environmental samples allows a "chemical fingerprint" to be generated through chemical analyses, which can be used to link contamination with a specific source [Wang and Stout, 2008]. Although, a more appropriate term may be "chemical signature", as environmental processes are likely to al-

ter contaminant composition over time, unlike a fingerprint which remains the same throughout a person's lifetime.

The process of chemical fingerprinting encompasses a range of analytical techniques, mainly involving gas chromatography, to allow the origin of a sample to be traced [Christensen et al., 2005]. Comprehensive chemical fingerprinting of a contaminated site allows accurate source identification and aids legal proceedings to ensure that polluters are held financially liable for any damage to the environment [Mudge, 2008].

Gas chromatography (GC) is a technique used to separate the components of a mixture based on specific chemical properties. GC analyses are frequently employed in environmental forensics due to the complex nature of most contaminant samples. Gas chromatography is generally used in combination with another analytical technique, generally in the role of a detector to identify the separated components. For example, a non-specific detector such as flame ionisation detection (GC-FID) can be employed to simply generate peaks corresponding to the separated components. On the other hand, coupling to a mass spectrometer (GC-MS) can provide additional information on the chemical structure of individual sample components. Detailed technical descriptions of all gas chromatographic instrumentation used during the course of this project are presented in Chapter 3.

The output of gas chromatography is presented in the form of a two-dimensional graph known as a chromatogram (or chromatograph); where the x-axis represents time and the y-axis represents the intensity of the detector signal. Individual constituents of a sample are represented by peaks in the detector signal. The chromatographic pattern can allow the composition of samples and possible sources to be compared visually.

### 2.3.1   Tiered Analytical Approach

Estimates suggest that over 5.5 million tonnes of oil were released into the environment by accidental spillages between 1970 and 2009 [ITOPF, 2009]; thus analysis and source identification of spilled oils has received much global attention. A number of tiered approaches to chemical fingerprinting have been proposed in the literature [Wang and Fingas, 1997, Daling et al., 2002, Douglas et al., 2007]. Over the years, standard procedures recommended by the US Environmental Protection

Agency [US EPA, 2011] and the American Society for Testing and Materials [ASTM, 1974] have been gradually modified to meet the requirements of specific incidents, such as the *Erika* oil spill along coastal France [Mazeas and Budzinski, 2002]. Generally, the main objectives are similar; to characterize individual compounds and classes present, to quantify hazardous compounds and to identify marker compounds capable of differentiating between sources [Alimi et al., 2003]. Wang and Fingas [1997] state that to achieve these objectives, the analytical procedure must be specific, selective and sensitive.

The generally accepted route begins with chemical fingerprinting by GC-FID in tier 1, to provide chromatograms for comparison by pattern recognition. However, GC-FID may produce very complex chromatograms containing numerous coeluting peaks, making it difficult to identify and quantify individual components [Christensen et al., 2005]. As a result, tier 2 employs GC-MS analyses to allow identification of specific compounds within the sample [Wang et al., 2007]. For example, a variety of source-specific compounds (biomarkers) and diagnostic ratios can be used to aid source identification. Finally, the last tier involves data processing and interpretation using a range of statistical methods. A simplified version of the tiered approach is given by the flowchart in Figure 2.1.

**Biomarkers and Diagnostic Ratios**     Biomarkers are compounds formerly from living organisms which are highly resistant to degradation and have been shown to provide useful information on the source of petroleum compounds [Wang et al., 2007]. For example, triterpanes and steranes are commonly targeted in spilled oils.

Diagnostic ratios may also prove useful; these are ratios of specific compounds which are capable of differentiating between contaminant sources. A chemical class often investigated using diagnostic ratios are the polycyclic aromatic hydrocarbons (PAHs). These compounds are commonly quantified in environmental samples due to their carcinogenic, teratogenic and mutagenic properties [Menzie et al., 1992]. The US EPA has classified sixteen PAHs as priority pollutants based on their toxicity; the chemical structures and names of these PAHs are given in Figure 2.2.

PAH diagnostic ratios are calculated either using PAH derivatives at differing alkylation levels or isomeric PAHs at the same alkylation level [Wang et al., 2002]. The generally accepted nomenclature for alkylated PAHs uses the parent PAH abbreviation followed by a number denoting the
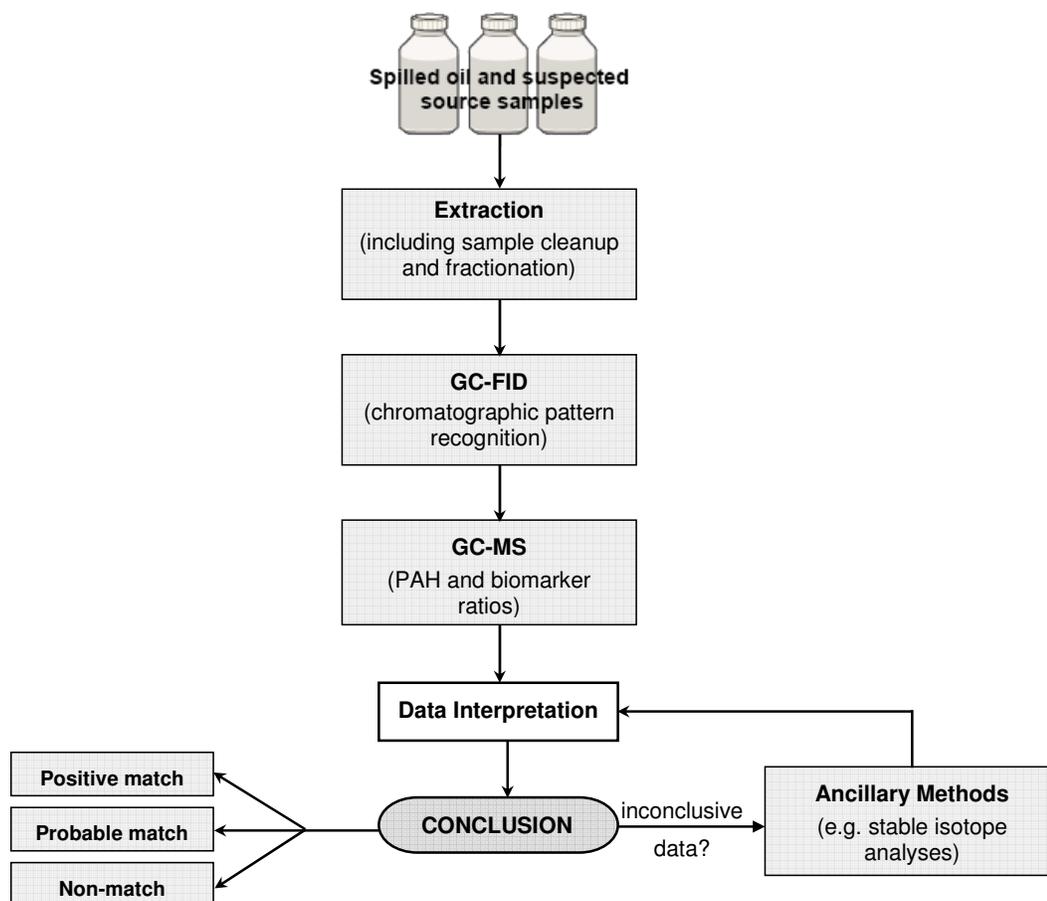
Figure 2.1: Flowchart indicating a general tiered approach for oil spill chemical fingerprinting (adapted from Wang et al., 1999).

alkylation level. For example, a naphthalene molecule with an alkyl sidechain containing two additional carbon atoms would be denoted as N2 (as shown in Figure 2.3). The distribution patterns of parent and alkylated PAHs can provide an overview of the method of PAH formation. PAHs can be produced by a variety of natural and anthropogenic sources. In nature, PAHs may be formed during biosynthetic reactions by plants or bacteria, volcanic activity or by combustion of natural matter, such as in forest fires. PAH sources produced by human activity include industrial processes, such as petroleum distillation and power generation, as well as various diesel/petrol engines and even cigarette smoke.



**Naphthalene (N)**  **Acenaphthylene (ACY)**  **Acenaphthene (ACE)**  **Fluorene (FLU)**

**Phenanthrene (PHE)**  **Anthracene (ANT)**  **Fluoranthene (FLT)**  **Pyrene (PYR)**

**Benzo[a]anthracene (BaA)**  **Chrysene (CHR)**  **Benzo[b]fluoranthene (BbF)**  **Benzo[k]fluoranthene (BkF)**

**Benzo[a]pyrene (BaP)**  **Dibenzo[a,h]anthracene (DBA)**  **Benzo[g,h,i]perylene (BP)**  **Indeno[1,2,3-cd]pyrene (IP)**
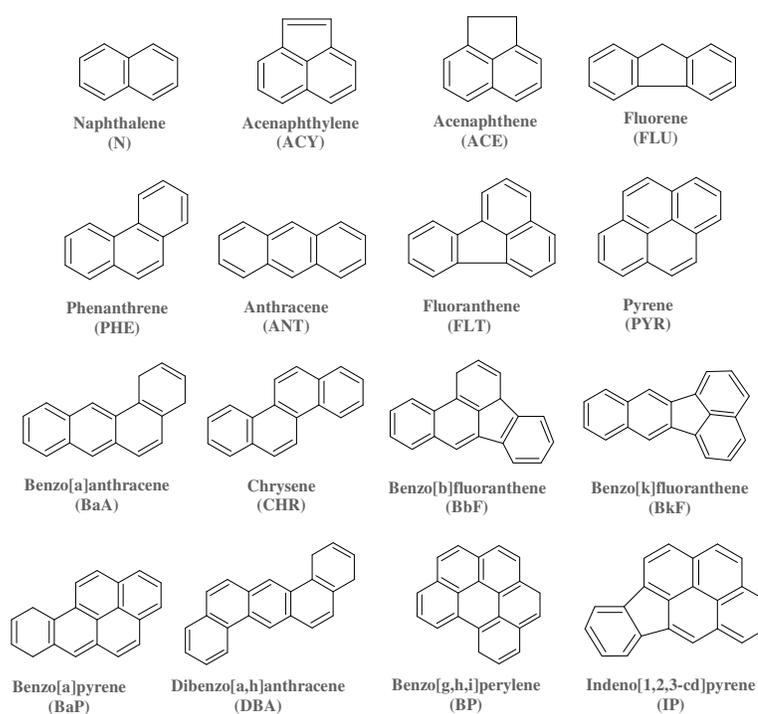
Figure 2.2: Priority pollutant polycyclic aromatic hydrocarbons (PAHs) as determined by the US EPA [USEPA, 2011b]. The terms in brackets indicate the commonly used abbreviations for each compound.
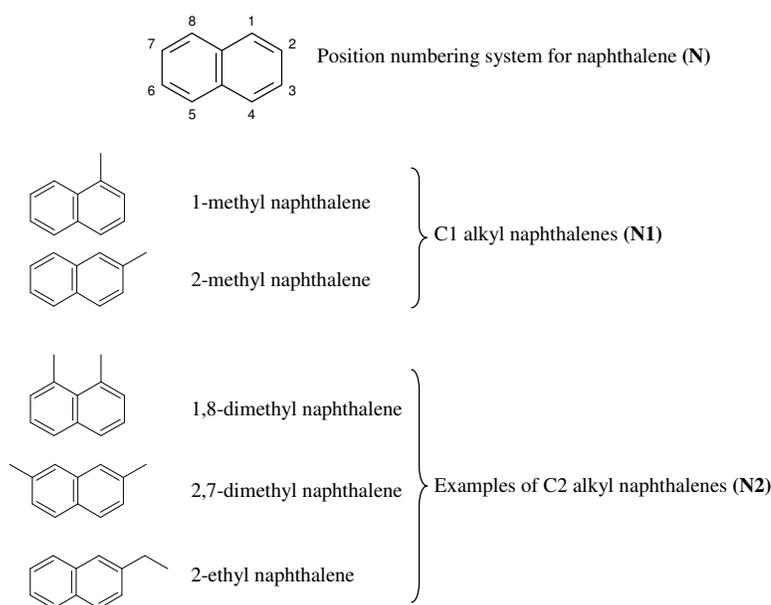
Figure 2.3: Examples of alkylated naphthalene chemical structures, including systematic numbering and common nomenclature.

Petrogenic PAHs are those resulting from petroleum sources, such as diesel, while pyrogenic PAHs are the products of combustion processes, such as the burning of fossil fuels. Pyrogenic sources display a characteristic PAH distribution where the parent PAH is dominant and alkyl PAH concentrations decrease with increasing alkylation level. The PAH profile of petrogenic samples displays the opposite pattern, where parent PAHs are in lowest abundance; both PAH profiles are displayed in Figure 2.4. Unlike biomarkers, certain PAHs may be affected by degradation processes. Consequently, the degree of weathering present in a sample may also be estimated using PAH diagnostic ratios [Wang et al., 2007].
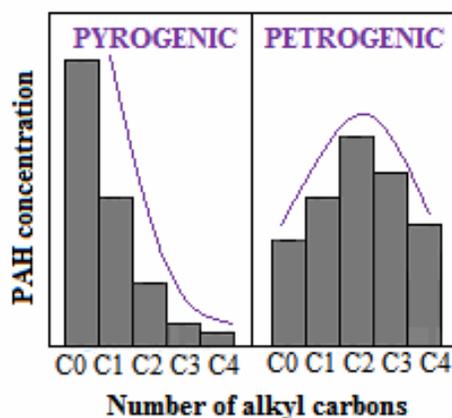
Figure 2.4: Illustration of typical petrogenic and pyrogenic PAH distribution patterns.

## 2.3.2   Applications in Environmental Forensics

As previously mentioned, the development of tiered analytical approaches was in direct response to major oil spill incidents, therefore there is a wealth of data in the literature regarding oil-oil and oil-source correlations.

Boehm et al. [1997] applied a modified tiered approach to the analysis of *Exxon Valdez* spilled oil samples two years after the incident. The authors showed that weathering had compromised the samples and rendered GC-FID data useless, thus chromatographic pattern recognition was not possible. In spite of this, Boehm et al. [1997] demonstrated that GC-MS data maintained the capacity to distinguish between oil sources. The use of a double diagnostic ratio plot (Figure 2.5) allowed the distinct separation of two oil sources; the *Exxon Valdez* oil and that from the Katalla oil field in the Gulf of Alaska.

The double ratio plot shown in Figure 2.5 demonstrates that C2 and C3 alkyl homologues of phenanthrene (labelled P2 and P3 respectively) and dibenzothiophene (D2 and D3 respectively) are weathered at approximately the same rate, thus can be coupled to provide weathering-resistant source ratios. Boehm et al. [1997] showed that two distinct sources were present in the sediment dataset; the subtidal sediments correlated with oil from the Katalla oil field in the Gulf of Alaska, while the oiled beach sediments were more similar to the *Exxon Valdez* tanker oil.

Wang et al. [2004] defensibly determined the source of a mystery oil spill in the Detroit River using a tiered approach combining GC-FID and GC-MS analyses. Two separate oil spills occurred
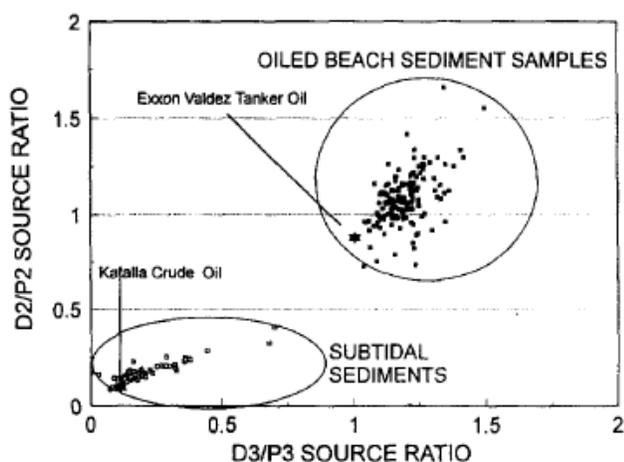
Figure 2.5: Weathering resistant double ratio plot of alkyl PAH homologues for oil spill source identification, where D2 and D3 are the C2 and C3 alkyl homologues of dibenzothiophene respectively, while P2 and P3 are the C2 and C3 alkyl phenanthrenes (copied from Boehm et al., 1997).

within a short time period in 2002 in the Rouge River, which leads into the Detroit River. Three separate locations were sampled along the rivers and resulted in similar chemical fingerprints. GC-FID and GC-MS analyses determined that the contamination was a mixture of lube oil with smaller quantities of diesel oil. Furthermore, using PAH distribution profiles, Wang et al. [2004] uncovered the presence of an additional pyrogenic PAH source relating to waste lube oil subjected to combustion/motor lubrication processes. This demonstrates the wealth of information which can be extracted from samples using simple analytical procedures and thorough data interpretation, allowing identification of the parties responsible for pollution.

The landfall of Hurricane Katrina in 2005 resulted in the release of thousands of gallons of crude oil from a damaged storage tank at a Louisiana oil refinery owned by Murphy Oil Inc. [Stout et al., 2007]. An extensive environmental forensic investigation was launched to determine the extent of the contamination. The widespread destruction caused by Hurricane Katrina resulted in numerous hydrocarbon sources at the sampled locations, including organic plant debris, peat and vehicle oil leaks. Therefore, the chemical fingerprinting method had to be selective for the Murphy crude oil. Stout et al. [2007] were able to prove the boundaries of the impacted zones by two-tiered (GC-FID and GC-MS) analysis of samples. Despite the degree of weathering and mixing present in the samples (enhanced due to the harsh weather conditions) evaluation of biomarker ratios using GC-MS allowed positive identification of samples containing Murphy oil (Figure 2.6). The study

by Stout et al. [2007] was conducted during the course of a class action law suit against Murphy Oil Inc. and the precise allocation of impacted zones allowed a settlement to be reached.

These literature examples illustrate that chemical fingerprinting is an established technique in oil spill investigations which can defensibly trace the origin of contamination.

### 2.3.3   Ancillary Techniques

However, in some cases, tiered approaches require modification, as they do not provide adequate data on the age and source of contamination [Uhler et al., 2002]. Therefore, the correlation between samples and suspected sources may prove inconclusive or ambiguous. The main problem with tiered approaches is the implementation of conventional GC techniques which do not have the capacity to resolve the complex composition of certain environmental samples [Gaines et al., 2007]. Advanced analytical techniques such as stable isotope analysis and two-dimensional gas chromatography have been developed with the aim of resolving these issues.

Stable isotope techniques measure the ratio of stable isotopes of an element present in an entire sample (bulk stable isotope analysis) or within a single molecule (compound specific isotope analysis). The main advantage of this technique is the ability to distinguish between chemically identical compounds based on their isotopic composition. Two-dimensional gas chromatography, on the other hand, is an advanced form of gas chromatography which provides separation of sample components based on two specific chemical properties instead of one.

Using these advanced techniques, this work aims to develop an ultra-resolution procedure for chemical fingerprinting in environmental forensic investigation of coal tar. Former manufactured gas plants are a ubiquitous source of contamination in the United Kingdom due to the by-product coal tar. Nevertheless, a standard procedure to adequately characterise coal tars has never been developed. The harsh pyrolysis conditions used in their formation, resulted in the degradation of any biomarkers from the parent coal, therefore, the chemical fingerprinting requirements vary from crude oils. The following sections will focus on the history of the manufactured gas industry and how the range of manufacturing processes may have affected coal tar composition.
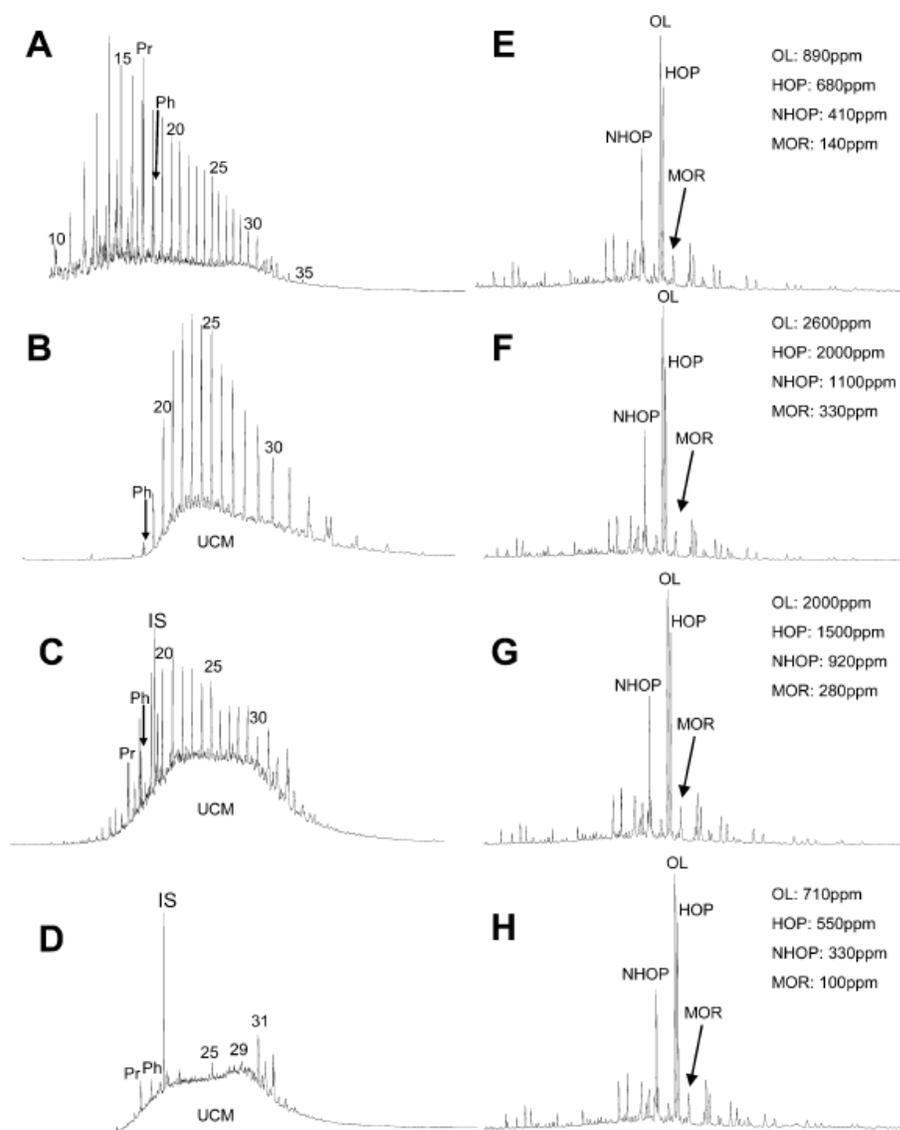
Figure 2.6: Chemical fingerprints of samples containing spilled crude oil. A-D represent tier 1 GC-FID chromatograms of samples containing crude oil in variable states of weathering and mixing while E-H show the corresponding tier 2 GC-MS partial ion chromatograms (m/z 191). Numbering shows the n-alkane carbon number; UCM=unresolved complex mixture; NHOP= 17r(H),21â(H)-norhopane; HOP= 17r(H),21â(H)-hopane; OL=oleanane; MOR= 17â(H),21r(H)-moretane; and IS= internal standards [copied from Stout et al., 2007].

## 2.4   History of Manufactured Gas

Manufactured gases are fuel gases derived from a solid or liquid source of organic carbon. Coal was carbonised by heating in an oxygen-depleted environment to produce coal gas, coke, tar and ammoniacal liquor [Murphy et al., 2005]. The principle of manufactured gas production for lighting and heating was developed by the British engineer William Murdoch in the late 18th century. Murdoch employed the first industrial-scale manufactured gas plant to light a Boulton and Watt Soho factory in 1798 [Findlay, 1917]. However, the first public gas works in Britain did not begin operation until 1813 [Thorsheim, 2002].

The manufactured gas industry expanded rapidly over the following 100 years, especially within Europe and the United States [Thomas and Lester, 1994]. By the mid 19th century, every town and city in Britain had at least one gasworks plant [Thorsheim, 2002]. The manufactured gas industry evolved over almost 200 years, with many advancements in technology providing better gas quality and quantity. However, the discovery of natural gas supplies in the mid 20th century resulted in the industry's rapid decline.

A concise list of former gas works sites has never been published, but it is estimated that approximately 3000±1000 former MGP sites exist within the United Kingdom alone [Thomas and Lester, 1994]. Figure 2.7 presents the known and suspected locations of FMGPs across the British Isles. These sites could prove harmful to the population and the environment, before, during or after redevelopment of the land if appropriate remediation techniques are not used [Thomas and Lester, 1994].

Manufactured gas was generally used for lighting and heating, however, a number of markets also developed for the consumption of the considerable quantities of by-products generated by the manufacturing process [Thomas and Lester, 1994]. During the manufacturing process, by-products, such as coal tars and ammoniacal liquors, and impurities, such as hydrogen sulphide and hydrogen cyanide, were removed from the gas [Hatheway, 2002]. Crude coal tar is a dark-coloured, non-aqueous phase liquid, or NAPL [Hatheway, 2006]. When coal tars are released into the environment the more soluble components migrate into the groundwater leaving the residual coal tar as a dense non-aqueous phase liquid (DNAPL). Consequently, these coal tar DNAPLs seep down through the water table until they reach a less permeable layer, at which point they begin to spread laterally to
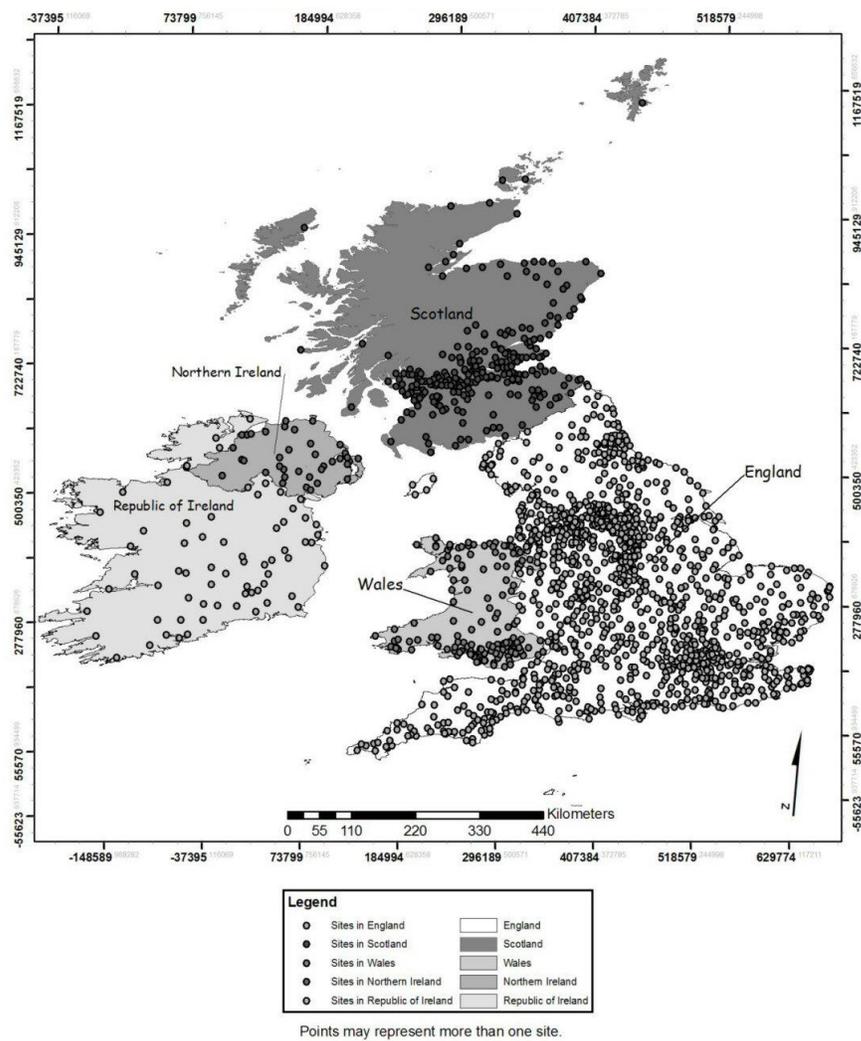
Figure 2.7: Known and suspected locations of former MGP sites in the British Isles (copied from Hatheway and Doyle, 2006).

Figure 2.8: An example of coal tar contamination at a former gasworks (copied from Wehrer et al., 2011).

form underground plumes [D'Affonseca et al., 2008]. An example photograph of coal tar DNAPL contamination leaching through the soil at a former gasworks is shown in Figure 2.8.

Coal tar DNAPLs are a common subsurface contaminant found at former manufactured gas plants (FMGPs). They are composed of thousands of organic and inorganic compounds, many of which may be found in trace quantities [Birak and Miller, 2009]. A large percentage of coal tar composition is made up of PAHs, including the sixteen priority pollutant PAHs previously shown in Figure 2.2. The vast majority of coal tar components are toxic, with many compounds known to have carcinogenic, teratogenic and mutagenic properties [Benhabib et al., 2010].

### 2.4.1   General manufacturing process

As previously mentioned, manufactured gas is a fuel gas derived from a source of organic carbon, such as coal or oil. In the early years of the manufactured gas industry, coal was the main carbon source utilised in Great Britain. The process conditions, such as temperature, carbonisation method and coal composition all played a role in determining the composition of the coal gas and any by-products formed [Brown et al., 2006]. The various production methods will be described in detail in later sections as it is imperative for environmental forensic investigators to have an understanding of the historical processes which resulted in a contaminated FMGP site. Firstly, the general

procedure of gas manufacture will be described, using the example of coal retort carbonisation, as this represented the earliest form of gas production.
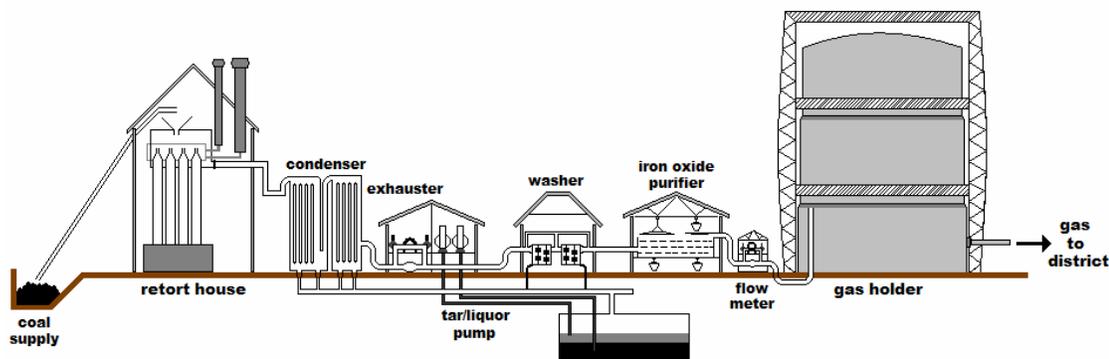


Figure 2.9: Schematic of the processes involved at manufactured gas plants (modified from Thomas and Lester, 1994).

Retort stands were the sealed, cast iron cylinders in which coal was placed. Coal was heated in these airtight retorts to allow volatiles to be driven off as a gas, which exited the retort via an exhauster pump [Hatheway, 2002]. A number of retort types were employed throughout the gas manufacture era, including horizontal, inclined and vertical designs, these will be described in further detail in later sections. The initial gas produced by heating coal contained a number of impurities which had to be removed prior to use for lighting or heating. Any moisture content present in the coal (particularly if it had been stored outdoors) created a fluid known as ammoniacal liquor. Purification processes were employed to remove this liquor, residual hydrocarbons, sulfur compounds and tar from the gas [Hatheway, 2002]. The final gas product distributed to consumers consisted mainly of hydrogen, carbon monoxide and a small portion of hydrocarbons [Tarr, 2004].

The main components of the gas processing portion of a gasworks site are described in the following paragraphs.

**Exhauster**   An exhauster was essentially a pump used to draw the gas stream out of the carbonisation chamber towards the purification processes. The exhauster would then pump the gas through the remaining sections of the gas plant process and into the gasholders [Hatheway, 2002]. The rate of gas flow, as determined by the exhauster, could cause a major effect on the by-products produced during carbonisation. If the coal gases were kept in contact with the hot retort walls

for longer periods of time, further degradation of organic compounds may have occurred resulting in differences in coal tar composition [Thomas, 2011a]. This is one of many important operating parameters which affected by-product composition and must be considered in present day environmental forensic investigations.

**Condensers**   The first stage in the purification process involved passing the gas stream through condensation towers to precipitate coal tars by lowering the temperature of the gas stream [Murphy et al., 2005]. Any coal tars and liquors condensed and drained into a tar and liquor tank (or well).

**Washers and Scrubbers**   The secondary purification steps involved washing the gas to remove any remaining tars and phenolic compounds. The washer typically involved bubbling the gas through fine perforations in the piping into an aqueous solution of ammoniacal liquor [Murphy et al., 2005]. This process dissolved any residual phenol and ammonia from the gas stream forming more concentrated ammoniacal liquor. The liquor would once again drain into the tar and liquor tank where it would separate by gravity and float on the surface of any tars present.

Following washing of the gas, a scrubber was employed to remove final traces of phenol and ammonia. Scrubbers consisted of tall, cylindrical towers filled with coke, bricks or wooden boards [Lewes, 1912]. The gas flowed up through the tower where it was sprayed with a mist of cooled water. This absorbed any residual ammonia and phenol, producing a weak ammoniacal liquor.

**Purifiers**   The final stage of the gas purification process involved the removal of poisonous compounds; namely hydrogen sulphide and hydrogen cyanide. Hydrogen sulphide was the most significant as the main combustion products are toxic sulfur dioxide and sulfuric acid which would degrade the interior of buildings [Thomas, 2011a]. Purifiers consisted of sealed boxes filled with lime (and later iron ore) to absorb the toxic compounds [Hatheway, 2002]. Eventually the purifying material became saturated with sulfur and had to be excavated from the boxes, resulting in spent oxides becoming an additional waste product of gas manufacture [Hatheway, 2002].

**Gasholders**   Lastly, the purified gas was passed through to a gasholder for storage. Typically, the gasholder could store 24 hours worth of maximum capacity gas production [Hatheway, 2002]. Gasholders were generally tall, cylindrical vessels which were sealed at the top but open at the base.

The base sat within a water-filled tank to prevent the escape of gas [Hamper, 2006]. As the gas supply filled the vessel it would rise out of the water tank. Initially, only single lift gasholders were available, but in 1924 the telescopic gasholder (with up to five lifts) was developed for additional storage space [Thomas, 2011a]. Regardless of the numerous purification processes, small portions of organic compounds remained in the final gas stored within gasholders. Consequently, the gasholder tanks would begin to fill with PAH-contaminated sludge [Hamper, 2006]. The area around former gasholders represents one of the many possible contamination hot zones at FMGP sites.

**Tar and Liquor Tanks**   Tar and liquor impurities were drained into storage tanks by gravity. Gravitational separation of liquors from the tars then occurred, resulting in a layer of ammoniacal liquor floating on the surface of coal tar. The tanks were usually underground and constructed from either cast iron plates, brick or clay. The design and size of the tanks were dependent on the size of the gas plant [Hatheway, 2002]. A pump was used to remove coal tar from the wells to above ground storage tanks ready for distribution as a feedstock for various industrial processes. The heat produced by the steam pump reduced the viscosity of the tar allowing it to flow easily from the tanks [Thomas, 2011a]. It can be assumed that these storage tanks would allow minor leakages over the years of operation, as well as accidental spills during coal tar transport and distribution. Consequently, the areas around former tar tanks at FMGP sites are potential contamination hot zones.

**Tar Distillation**   In 1901, the worldwide production of coal tar was estimated at over 2.6 million tonnes, with the UK providing approximately 34% of the total [Findlay, 1917]. By 1910, the annual production of coal tar in the UK exceeded 1.3 million tonnes [Findlay, 1917]. In the early years of gas production, coal tar was thought to be an undesirable waste product which was extremely difficult to dispose of. It was often burned as a fuel, until 1838, when it was discovered that coal tar could provide useful products such as creosote oil for preservation of timber [Findlay, 1917]. Furthermore, in 1945, research by Professor Hoffman at the Royal College of Chemistry in London demonstrated the isolation of principal coal tar constituents [Findlay, 1917]. This discovery led to coal tar becoming the base feedstock for production of chemicals prior to the development of the petroleum industry. A multitude of chemical products were made using coal tar, including dyes, explosives, drugs and antiseptics [Murphy et al., 2005]. These additional uses of coal tar mean it is a possible contaminant at sites other than FMGPs.

## 2.5    Evolution of Manufacturing Processes

The early philosophy of coal gas, as proposed by William Murdock in 1792, was the use of small-scale gasworks to supply individual establishments. The industry, of course, was expanded to include large-scale, commercial gasworks using gas mains to distribute the gas supply to whole towns and even cities. The technology applied at these sites evolved to produce a higher quality and quantity of gas. Thus, it is imperative that environmental forensic scientists have knowledge of the various processes which may have occurred at MGP sites as this will have affected the resulting contamination. The evolution of these manufacturing processes will be described in the following sections.

### 2.5.1    Low Temperature Horizontal Retorts

Remote or rural buildings could not connect to gas mains, thus they required an alternative gas source. Small country gas works were consequently used to illuminate such residences and were designed in such a way as to enable a single person to operate the entire process. The retort house would generally contain a single retort bench, consisting of three retorts and a furnace, where coal charging and coke removal would be performed manually. Due to the small scale of such gas works, very simple gas technology was generally used throughout the years of operation. For example, low temperature horizontal retorts were initially used in all early gasworks, as they were the first retort design to be developed. The technology at large scale gasworks evolved to improve gas quality and quantity produced, while small scale gas works were limited to these low temperature horizontal retorts.

The retorts were heated directly by an underlying dedicated furnace. The process was deemed low temperature as a directly-fired retort may only reach temperatures of around 600 °C [Harkins et al., 1988]. This resulted in low gas production and limited the degradation of organic compounds. Therefore, any by-product coal tars produced by this method should exhibit a lower range of organic compounds than higher temperature retorts and be similar in nature to the parent coals [Soule, 1922].
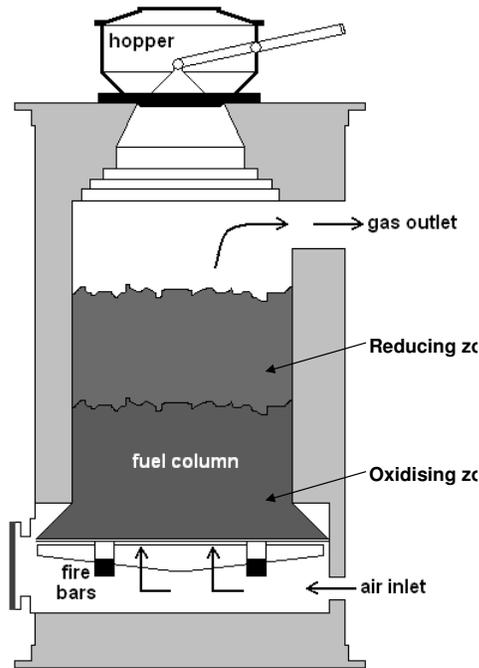
Figure 2.10: Schematic of a producer gas reaction vessel (adapted from Thomas, 2011a).

## 2.5.2   Producer Gas

Producer gas became popular in the late 1800s and involved the gasification of coal. Unlike coal carbonisation, where coal is heated under oxygen-depleted conditions, gasification of coal is the partial combustion in air/oxygen to convert the entire feedstock into combustible gases [Francis and Peters, 1980].

The producer gas process involved drawing air through a thick, heated bed of fuel (coal or coke) within a closed reaction vessel [Francis and Peters, 1980]. A schematic of the producer gas process is given in Figure 2.10. The fuel bed was heated using fire bars or grates at the base of the fuel column. As the fuel source depleted, additional coke or coal was placed in the vessel via a hopper. The air stream was drawn in through the base of the vessel and gradually passed up through the fuel column and exited the vessel as producer gas. The depth of the fuel bed created two well-defined zones, an oxidising zone and a reducing zone, as represented by the reaction given in Equation 2.5.1 [Francis and Peters, 1980].

$$C + O_2 \rightarrow CO_2 \, (\textit{oxidation zone}) \rightarrow CO_2 + C \rightleftharpoons 2CO \, (\textit{reducing zone}) \tag{2.5.1}$$
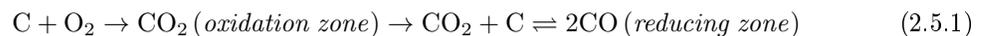
Carbon dioxide is first formed in a thin oxidation zone at high temperatures (>1200 °C) in an area close to the heating source. This is then reduced to carbon monoxide by further reaction with the carbon source. The final producer gas was generally a mixture of carbon dioxide, carbon monoxide, hydrogen, nitrogen and gaseous hydrocarbons [Francis and Peters, 1980]. If coke was used as the carbon source, no significant amounts of tar formation would be expected [Francis and Peters, 1980]. However, specific types of producer gas, such as Mond gas, were designed to use bituminous coal as the carbon source, thus would produce quantities of coal tar. Producer gas was rarely used for distribution to consumers, it was generally manufactured to heat other processes, such as high temperature retort carbonisation.

## 2.5.3   High Temperature Horizontal Retorts

Retort technology advanced during the 19th and 20th century to allow gas production of a better quality and quantity. Around 1881, semi-gaseous and gaseous fired retorts were introduced to replace the direct-fired low temperature retorts [Williams, 1877]. In the semi-gaseous form, a fuel bed below the retort was still employed, however, unlike the early retorts the air flow to the fuel bed was restricted to allow production of carbon monoxide. Additional air flow was then introduced to burn the carbon monoxide, generating higher temperatures [King, 1948].

The gaseous fired retort was a further development which used a gas producer to heat the retorts [Butterfield, 1904]. The producer gas was circulated around and between the retorts, where it was burned with a secondary air supply. The resulting hot exhaust gas was channelled around the retort flues to heat the retort coal [King, 1948]. These advanced retort settings were expensive to construct [Thorsheim, 2002], however, they were more fuel efficient and had greater temperature control, capable of exceeding 1000 °C [Butterfield, 1904].

The main disadvantage of horizontal retorts was the intermittent nature of the process. During carbonisation the coal within the retorts was completely converted to coke which required removal [Thorsheim, 2002]. Initially, coke removal was performed manually, until a mechanical arm system was developed for charging retorts with fresh coal and removing coke [Butterfield, 1904]. However, even with the mechanical arm, the gasification process had to be halted, coke removed and the retort replenished with a fresh supply of coal before the process could begin again [Butterfield, 1904].

### 2.5.4   Continuous Vertical Retorts

The evolution of inclined retorts was aimed to resolve the issues associated with the horizontal retort design. The goal was to make retort loading and unloading processes easier by gravity assistance due to the 30° inclined angle. However, it was difficult to get an even coal charge at the inclined angle and coke often jammed within the retort [Harkins et al., 1988].

Continuous vertical retorts were introduced in the early 1900s with a view to solving the problems with inclined retorts [Hamper, 2006]. The process was highly automated, therefore suitable for large-scale gas production. The 90° retorts were charged with a deep layer of coal via a conveyor belt attached to a hopper at the top of the retort (Figure 2.11). The coal passed by gravity through the retort and was gradually carbonised, before exiting the base of the retort as coke. An extractor, operating an Archimedes screw method, allowed continual removal of coke into waiting coke cars [Thomas, 2011a]. All vertical retort designs employed a producer gas system for heating, as in high temperature horizontal retorts.

The vertical retorts provided continual gas production, thus gave better efficiency than horizontal retorts. However, smaller gasworks tended to retain low temperature horizontal retorts as it was not economical to upgrade the technology [Thorsheim, 2002]. On the other hand, large scale gasworks may also have retained horizontal retorts as they produced higher quality, aromatic-rich tars which were more valuable as a chemical feedstock for the production of dyes and explosives [Findlay, 1917].
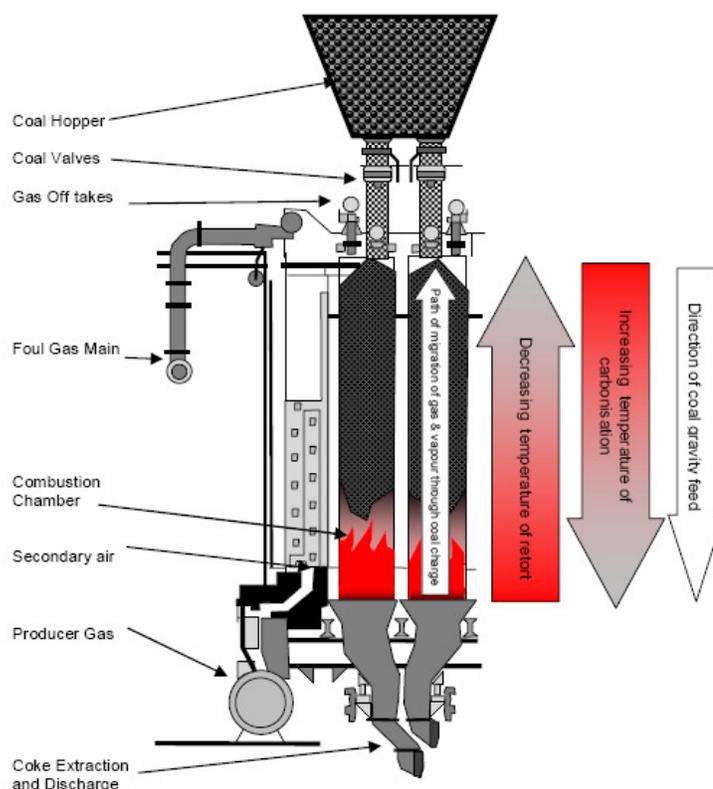
Figure 2.11: Schematic of a continuous vertical retort used for coal gas manufacture (copied from Thomas, 2011a).

## 2.5.5   Carbureted Water Gas

As the gas industry thrived, market demand increased and existing retort gas plants struggled to satisfy these demands. Water gas plants were introduced to Britain in the late 1800s in an attempt to resolve this problem. The plants required less space and capital outlay than early gasworks designs, yet could increase gas production in a matter of hours to meet peak demands, while retort gasworks required at least three days to increase production [Tarr, 2004].

In this process, steam was passed over a carbon source (usually heated coke) to yield gas. The reaction was extremely endothermic, meaning it absorbed heat energy, therefore water gas was generally produced by a two-step cyclic process referred to as the "run" and "blow" cycle [Harkins et al., 1988]. The plants alternated between the blow stage, where coke was heated to incandescence by an air stream, and the "run" stage, where steam was injected into the coke. During the "run"

27

step, the fuel would gradually cool due to the endothermic nature of the reaction, therefore the cycle would be repeated to reheat the vessel.

The main products of the reaction were hydrogen, carbon monoxide and carbon dioxide. To improve the calorific value of the gas, an oil spray was introduced to the hot gas stream within a superheater. This process induced thermal cracking of the oil constituents into gaseous compounds [Murphy et al., 2005]. The enriched gas was deemed "carbureted" water gas, or CWG for short. Water gas without carburetion tended to produce insignificant levels of by-product tar [Hamper, 2006].

In the United Kingdom, CWG plants consumed mainly light oils for carburetion [Butterfield, 1904], while in the USA heavier or crude oils tended to be used, especially after World War I when light oils such as gasoline were in high demand [Harkins et al., 1988]. However, the type of oil was generally dependent on their availability and cost and may have changed frequently [Harkins et al., 1988]. Heavy oils produced tars with a high water content, often referred to as tar-water-emulsions. Tar separators, usually working by gravitational separation, were employed to separate tars from gas liquors. In the USA, unlined tar lagoons or tar ponds were often used to separate out the by-product phases and as such can pose a major environmental hazard today [Hatheway, 2002]. The composition of the tar emulsions was mainly dependent on the carbon source and type of oil used for carburetion.

## 2.5.6   Coke Ovens

Coke ovens were mainly located at iron and steel works where coke was required for the smelting process and can still be found in operation at present-day steel works. A small number were employed at gasworks for gas production, although the gas may also be supplied directly as a by-product from coking works and supplemented with coal gas produced by other gas-making processes during periods of peak demand [Gardner, 1915]. Initially, coke ovens were not used with the intention of recovering by-products such as coal tar. It was only when the value of coal tar as a chemical feedstock was discovered that "by-product" coke ovens became popular.

The theory of operation of coke ovens is similar to that of retort stands, however, coal was generally crushed prior to placement in a coke oven for optimal coke formation. The coke ovens were stoked with crushed coal, ensuring a void space for gas evolution was left at the top of the oven, then

carbonised for prolonged periods [Butterfield, 1904]. The ovens could be heated by circulating either the gas produced by other coke ovens or by a separate producer gas plant. The former is generally used in present day coke ovens.

### 2.5.7   Gas and Oil Reforming Plants

Coal was not the only solid feedstock used in gas production, processes carbonising wood and peat were also developed [Francis and Peters, 1980]. However, the most popular alternative feedstock to coal was a liquid fuel. Prior to the development of the petroleum industry, derivatives of crude oil were relatively inexpensive. Consequently, many gasworks devoted a portion of their production process to oil gas.

Oils were converted to gas through thermal cracking in the presence of steam to produce fuel gas, alongside coke or carbon black and tar as by-products. Light oils produced low volumes of tar compared to heavy oils which gave denser tars with a high aromatic content. The main advantage of oil gas was the lack of ammonium and cyanide production which reduced the costs associated with gas purification [Harkins et al., 1988].

Cyclic catalytic processes were often employed in the UK for oil gas production. The plants were operated in a cyclic process similar to the "run" and "blow" method employed in CWG plants [Harkins et al., 1988]. In early reforming plants, thermal cracking in the presence of steam was performed at approximately 1000 °C. A lime or nickel catalyst was used to encourage reaction between the steam and oil hydrocarbons, increasing gas yield and reducing by-product formation [Department of the Environment, 1995]. Examples of oil gas processes used in the UK were the SEGAS and micro-simplex methods. The micro-simplex process was used to reform hydrocarbon gases and liquefied petroleum gases, while the SEGAS process was capable of gasifying a range of oil feedstocks, including residue oils. The oil reforming process was relatively clean, with only minute quantities of tar produced.

### 2.5.8   Complete Gasification

Complete gasification plants were less popular than other discussed methods and combined the use of vertical retort coal carbonisation with water gas processes. Coke produced within the vertical

retort was transported to a water gas plant which could operate with or without carburetion. The complete gasification process was used within Tully gas plants in the UK. Tully plants were unreliable and difficult to operate, thus were often used solely as water gas plants [Thomas, 2011a].

## 2.6   Factors Effecting Coal Tar Composition

The composition of tar wastes at FMGP sites can vary across different sites and even within the same site, due to the variety of manufacturing methods and process parameters which may have been altered. It is important for environmental forensic investigators to consider these factors during the analysis and evaluation of contaminated samples from FMGP sites.

### 2.6.1   Coal Type and Particle Size

As expected, the type of raw material used in gas manufacture was a major influence on the by-product coal tars, therefore a basic understanding of coal pyrolysis is required to allow any correlations with tar formation to be shown.

Veras et al. [2002] demonstrated the degradation of coal to produce volatile gases and tars (Figure 2.12). This emphasises the effect of coal structure on coal tar composition. Pyrolytic conditions provide initial cleavage of alkane or ether linkages within the coal framework thus yielding alkyl PAHs and phenolic fragments [Metcalf, 1983]. Secondary reactions, including condensation, ring fusion or ring closure reactions, can then provide a further range of compounds.
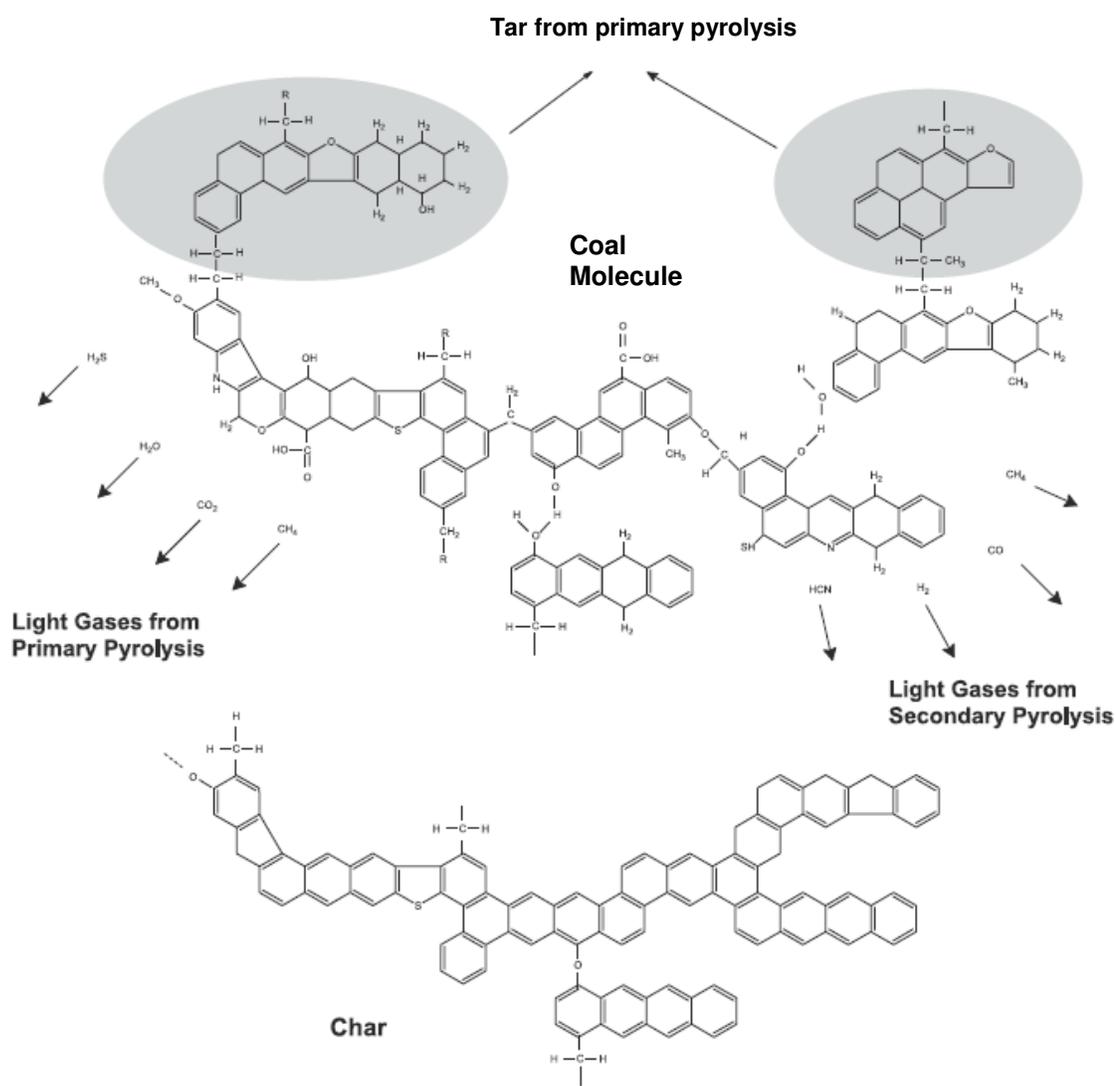
Figure 2.12: Model for the pyrolysis of coal (copied from Veras et al., 2002).

It has previously been noted that low temperature tars are more representative of the parent coals than tars produced by high temperature pyrolysis [Ibarra et al., 1989]. Veras et al. [2002] demonstrated a further correlation between coal rank and primary tar composition; higher ranking tars produced tars with a higher aromatic content. This coincides with a higher degree of aromaticity in the parent coal. This statement was supported in a study by Casal et al. [2008] which showed that at pyrolyis temperatures of approximately 500 °C there was a strong correlation between parent coal and tar.

Table 2.1: Comparison of tar yield produced by the carbonisation of coals from two different mains using both vertical and horizontal retorts and with varying coal particle size (data from Terrace, 1948).

| Coal tar (yields per ton of coal)[a] | Coal type | | | | | |
|---|---|---|---|---|---|---|
| | Mitchell Main, Wombwell, S. Yorkshire | | | Dalton Main, Rotherham, S. Yorkshire | | |
| | Large | Medium | Small | Large | Medium | Small |
| Horizontal gas retort | | | | | | |
| Coke (cwt) | 14.59 | 14.50 | 14.57 | 14.08 | 14.23 | 14.22 |
| Tar (gal) | 13.00 | 10.90 | 11.90 | 14.80 | 12.50 | 11.70 |
| Liquor (gal) | 4.9 | 5.50 | 5.20 | 8.10 | 7.40 | 7.30 |
| Gas (cf) | 13,460 | 13,280 | 13,440 | 13,060 | 13,055 | 13,090 |
| Continuous Vertical Retort | | | | | | |
| Coke (cwt) | 13.28 | 13.80 | 13.69 | 12.47 | 12.86 | 13.20 |
| Tar (gal) | 16.30 | 12.30 | 11.10 | 19.50 | 13.90 | 13.20 |
| Liquor (gal) | 10.50 | 16.20 | 17.30 | 14.80 | 14.00 | 13.10 |
| Gas (cf) | 21,495 | 19,470 | 17,730 | 21,265 | 20,175 | 19,620 |

[a]cwt = centum weight (1 cwt = approximately 50.8 kg), gal=gallon, cf=cubic feet.

Terrace [1948] noted an additional trend associated with coal particle size. In general, the carbonisation of larger coal particles produced larger yields of tar; the results of the study depicting the effect of coal type, size and retort type on gas production and tar yields are summarised in Table 2.1. It can be seen that larger coal particles from both the Mitchell and Dalton coal mains result in larger tar yields regardless of the manufacturing process employed; yet the effect appears exaggerated in continuous vertical retorts [Terrace, 1948]. This phenomenon may explain the lower quantities of tar produced by coke ovens, as coal was generally crushed prior to placement in the coke ovens.

## 2.6.2   Temperature

The temperature of carbonisation was also investigated by gas engineers as a possible way of improving gas production yields. The effect on coal tar composition was not an issue at the time as it was treated as a by-product, the most important issues would have been the quality and quantity of manufactured gas.

A study by Pryde [1934] reported that temperature variation could cause an effect on the physical characteristics of coal tar as well as chemical composition. Low temperature tar was reported as an

oily, brown liquid containing high quantities of unsaturated hydrocarbons, paraffins and phenols. The only true low temperature tars are those produced by early directly-fired horizontal retorts. On the other hand, high temperature tar was described as a viscose, black liquid containing only trace amounts of aliphatics but high quantities of aromatics [Pryde, 1934].

Rhodes [1945] further demonstrated the relationship between carbonisation temperature and tar composition (Figure 2.13). The research showed that low temperature tars are rich in acids and aliphatics, while high temperature tars have a dominant abundance of aromatics.
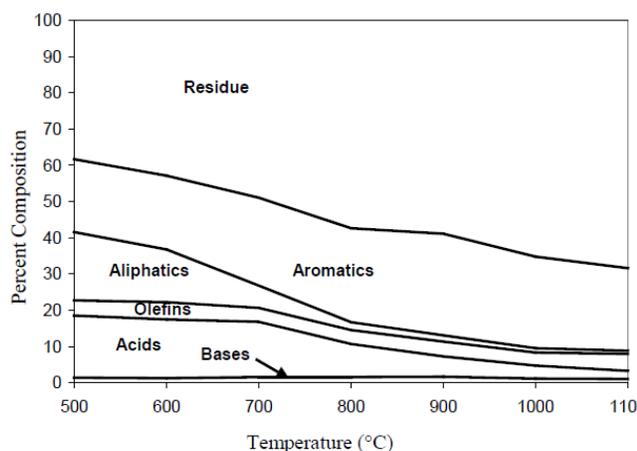


Figure 2.13: Relationship between temperature and the composition of coal tar. Copied from Emsbo Mattingly and Boehm [2003], originally from Rhodes [1945].

### 2.6.3   Manufacturing Process

**Retort Shape**   As discussed in the review of gas manufacturing processes, coal retorts were available in a number of designs, including horizontal, inclined and vertical structures. Inclined retorts did not develop much popularity and were extremely rare in the UK therefore this section will focus on the horizontal and vertical designs [Thomas, 2011a].

The tars produced by vertical and horizontal retorts varied due to the differences in retort shape. Young [1922] proposed that the duration of contact time between the evolved gases and the hot retort surfaces had a profound effect on the tar by-products. Contact with the heated retort walls resulted in additional thermal cracking of the constituent molecules and thus different coal tar components. The main parameter controlling contact time was the retort shape. The larger

exposed surface area of horizontal retorts provided a greater opportunity for contact of the gases with a heated surface and a greater chance of thermal cracking [Young, 1922].

A further parameter influencing the residence time of the gas within retorts was the gas flow. This was largely controlled by an exhauster which pulled gas from the retorts into the purification sections of the plant. The faster the rate of gas flow from the retort, the lower the chance of additional thermal cracking. Moreover, the chance of thermal cracking was reduced further if the gas was cooled quickly by condensers as it left the retorts [Thomas, 2011a].

Within CVRs, the deep coal charge had a continual downward motion as it was carbonised and converted to coke. Fresh coal near the top and central areas of the retort (i.e. furthest from the heat source) would have a lower temperature than coal near the base and hot walls of the retort [Young, 1922]. Consequently, the coal tar produced was expected to be of low temperature character, except for traces of high temperature tar from the coal close to the retort walls. Additionally, any gases evolving from the centre of the retort could rise vertically from the coal charge, avoiding contact with the retort walls and any chance of thermal cracking.

The depth of coal charge in horizontal retorts was obviously much lower than in vertical designs. In horizontal retorts, any variations in coal depth would affect the coal tar composition greatly due to the large effect on exposed retort surface area and the variation in void space above the fuel bed [Rhodes, 1945]. This effect was negligable in vertical retorts. A number of early publications have proposed the differences between vertical and horizontal retort coal tars in terms of the percentage composition of oil fractions [Young, 1922, Malatesta, 1920, Terrace, 1948]. All authors agree that horizontal retort tars contain higher abundances of aromatic hydrocarbons (especially naphthalene) and phenol, but less phenol homologues than those produced by vertical retorts.

**Carbureted Water Gas Tar**

Water gas plants produced insignificant quantities of tar, thus their composition will not be discussed in detail [Harkins et al., 1988]. On the other hand, carburetion of the gas produced tar with different characteristics to coal carbonisation tar. Meade [1934] reported that CWG tar has low viscosity/density compared to the equivalent formed by carbonisation. This caused additional disposal problems as the tar could easily form emulsions with water and required separation in

a settling tank before the tar portion could be removed for distillation [Hatheway, 2002]. CWG tar was noted to lack phenols and naphthalene, as well as a low content of nitrogen heterocycles, reflecting the low nitrogen content of coke and petroleum oils. However, CWG tars are reported as having a high paraffinic content due to carburetion [Harkins et al., 1988]. The tar composition could vary greatly according to the type of carburetion oil used, as heavy oil produced tars with higher viscosity and aromaticity than those produced using light oils [Harkins et al., 1988].

**Coke oven tar**

Coke ovens and producer gas plants provided high viscosity tars with similar characteristics. Coke ovens could be operated at a variety of temperatures; low temperatures ($<700$°C) gave tars similar to low temperature retorts, while high temperatures ($>700$°C) produced tars with high PAH content [Hamper, 2006]. However, the nitrogen-containing heterocycles were generally degraded further into ammonia, cyanide and nitrogen gas due to the long residence time (approximately 16 hours) within the coke ovens [Lewes, 1912]. High concentrations of PAHs were produced due to these longer residence times, as the hot gases within the oven were subjected to further degradation and ring condensation, such as in lightly charged horizontal retorts [Gardner, 1915]. Rhodes [1945] suggested that by-product coke ovens were capable of degrading gases to such an extent that aliphatics, tar acids and naphthenes were highly depleted in the resulting tar.

In conclusion, a vast number of factors can affect the final tar composition. The combination of these factors and the long periods of exposure (usually decades or even centuries) in the subsurface make coal tar characterisation extremely challenging. For simplicity, the major factors which influenced tar formation and the process-specific differences in tar composition are summarised in Tables 2.2 and 2.3 respectively.

Table 2.2: Summary of the primary factors influencing tar composition.

| Process | Factor/Parameter | Effect on Tar Formation |
|---|---|---|
| Coal Retorts | retort shape | larger exposed surface area of hot retort walls gives further degradation of gas components |
| | coal charge depth | mainly evident in horizontals as effect of depth on exposed retort surface is small in vertical retorts |
| | gas flow | slower flow rate gives greater opportunity for gas degradation |
| Coke Ovens | residence time | longer residence time allows further thermal cracking |
| CWG | oil type (for carburetion) | heavy oils provide higher aromaticity while light oils give low quantity of tar |
| | carburetion temperature | controls extent of thermal cracking of oil spray |
| All processes | coal type | composition of original raw feedstock |
| | coal particle size | larger coal size produces more tar |
| | temperature | lower temperatures give tars similar to parent coal composition |

Table 2.3: Comparison of tar characteristics when produced by various manufacturing processes. Compiled using data from Soule [1922], Warne [1907], Lewes [1912], Harkins et al. [1988].

| Gas Manufacturing Process | Effect on Tar Characteristics |
|---|---|
| Low Temperature Horizontal Retorts | brown, oily, low viscosity tar; high content of low molecular weight volatiles e.g. alkanes, alkenes, phenols, but low carbazole content |
| High Temperature Horizontal Retorts | black, high viscosity tar; rich in aromatics (e.g. naphthalene), high toluene and carbazole, but low creosote/anthracene fraction and trace aliphatic content |
| Vertical Retorts | oily, brown, low viscosity tars; lower naphthalene content than horizontal retorts, but higher abundance of tar acids |
| Carbureted Water Gas | very low naphthalene and phenols content, aliphatics always present |
| Coke Ovens | high PAH content, low in nitrogen heterocycles, acids and aliphatics |
| Producer Gas | very high viscosity; lacking in tar acids, aliphatics and lighter oils |

## 2.6.4   Environmental Risk

In the mid-twentieth century, the discovery of plentiful natural gas fields in the North Sea led to decommissioning and redevelopment of MGP sites within the British Isles [Harkins et al., 1988].

Likewise in the USA, the development of long-distance pipelines from natural gas fields negated the requirement for large-scale manufactured gas plants and led to the industry's decline [Hamper, 2006]. The remediation and redevelopment of former MGP sites is highly desirable as it increases the land value and eliminates any environmental liability. However, appropriate remediation strategies have only recently been implemented, such as the National Brownfield Strategy for the redevelopment of sites affected by previous land uses [Department of Communities and Local Government, 2008]. The National Brownfield Strategy estimates that approximately three quarters of new development occurs on brownfield land and provides a summary on how to prepare, remediate and reuse the land. In the past, adequate remediation measures were not always implented, for example, Bozek [1995] notes that MGP structures were often demolished and used to fill in tar wells, with clean fill layered over the site prior to redevelopment.

Coal tars were stored in tar wells or tanks at all FMGP sites and often sold on as a feedstock for the chemical industry [Francis and Peters, 1980]. Coal tars are exceptionally dense and contain high molecular weight compounds which are unlikely to degrade quickly when released into the environment. Consequently, tar wells and gasholders are the main sources of environmental concern at decommissioned sites, due to the high potential for residual tar contamination, which may pollute groundwater systems and cause a vapour risk for any overlying buildings [Hatheway, 2002]. Although the gas industry was generally well-managed, accidental spillages and leaks over the years of operation may have resulted in the presence of high concentrations of volatile organic compounds, semi-volatile organic compounds, PAHs, heavy metals and cyanide at FMGP sites [Gibert et al., 2007]. Many of these compounds are hazardous, especially in high concentrations. For example, the majority of PAHs exhibit high toxicity and carcinogenic properties, thus they pose a real threat to health and the environment if appropriate remediation strategies are not performed [Menzie et al., 1992, Cave et al., 2010].

The remediation of coal tar DNAPLs is a challenging process, as their chemical composition has been shown to vary dramatically within a single FMGP site, as well as between multiple sites [Brown et al., 2006]. Furthermore, the wide range of chemical, physical and toxicological properties of individual components in the coal tar requires extensive chemical fingerprinting. As previously mentioned, the US EPA has identified 16 priority pollutant PAHs (Figure 2.2), but in reality coal tars may contain between 500 and 3000 different PAHs and alkyl PAH isomers [Hatheway, 2002]. These compounds can also be found in crude oil, where chemical fingerprinting approaches have

evolved to provide accurate oil-oil and oil-source correlations [Wang and Stout, 2007]. However, the lack of biomarkers in coal tar requires the development of specific analytical techniques for their characterisation.

Accurate chemical fingerprinting is required at FMGP sites (and in the case of spilled oils) to ensure multiple sources of contamination are not present [Wang and Stout, 2007]. This could help to distinguish more modern spills from older gasworks contamination. Furthermore, for FMGPs split into multiple land holdings, accurate chemical fingerprinting can allow characterisation of contamination at various parts of the site and help to identify liability. Given the large number of former gasworks sites in the UK and the new "polluters pay" legislation, it is reasonable to assume there will be many liability cases in the future, thus an accurate method of chemical fingerprinting for source identification of coal tar is desperately needed.

Regardless of improvements in analytical techniques, there have been few reports on coal tar composition in the literature. Generally, the literature focuses on the challenges involved in characterisation of coal tar contaminated land for remediation purposes [Hatheway, 2002, Birak and Miller, 2009]. For example, Birak and Miller [2009] state that full characterisation of coal tars at FMGP sites is still limited by analytical techniques, reiterating the gap in scientific knowledge which the chemical fingerprinting method developed in this project can help to fill. The following sections will review previous research on the chemical composition of coal tar, with emphasis on data useful for environmental forensic investigations.

## 2.7   Chemical Analysis of Coal Tar

An extensive volume of literature is available on the manufactured gas industry and by-products, however, the majority of this data was not produced using modern gas chromatographic technology. Nevertheless, information on coal tar composition and comparisons between manufacturing processes were still reported which could aid studies in the present day.

Early investigations into coal tar composition focussed on the percentage abundance of each distilled fraction within the total sample, as analytical techniques limited the amount of specific chemical information that could be obtained. Nevertheless, a dissertation by Soule [1922] demonstrates the differences in coal tar fractions produced by different manufacturing processes; Table 2.4 summarises

Table 2.4: Comparison of boiling point fractions of tars produced by various manufacturing processes (data taken from Soule, 1922).

| Boiling point range | Percentage of total weight | | |
|---|---|---|---|
| of fraction (°C) | Low Temperature Tar | Coke Oven Tar | Gasworks Tar |
| 20-173 | 0.64 | 0.70 | 1.19 |
| 173-237 | 9.19 | 8.27 | 9.22 |
| 237-281 | 12.50 | 12.44 | 10.12 |
| 281-315 | 6.94 | 6.21 | 3.78 |
| 315-326 | 2.90 | 2.53 | 1.93 |
| Pitch (difference) | 67.83 | 69.85 | 73.76 |
| Melting point of pitch (°C) | 53 | 69 | 89 |

the main findings. The lack of specific and sensitive analytical methods during this time meant that specific chemical variation between the fractions were not reported.

A study by Warne [1907] also noted that the increased pitch and high molecular weight content of high temperature gasworks tars resulted in a denser tar compared to those of low temperature processes. This provided a simple, initial method of tar characterisation by visual comparison.

The development of gas chromatography in the 1950s allowed a more thorough examination of coal tar constituents. Novotny et al. [1981] reported the development of a GC-MS method to identify specific compounds in coal tar; namely PAHs, alkyl PAHs and acidic compounds (through prior derivatisation). The method employed a rigorous seven-step liquid-liquid extraction and fractionation scheme, which at the time was considered a fast and efficient method of sample preparation. Additionally, low instrumental sensitivity resulted in the identification of only major components, trace compounds could not be analysed. Novotny et al. [1981] noted that the harsh environmental impact of coal tar, made studies on such contamination highly desirable.

During the 1980s, numerous studies focussed on the identification of PAHs in coal tar [Romanowski et al., 1983, Later et al., 1981, Tomkins et al., 1980]. Burchill et al. [1983] defied this trend by analysing the nitrogen-containing compounds in coal tar using GC coupled to a nitrogen-selective alkali flame detector. Burchill et al. [1983] identified approximately forty nitrogen heterocycles within each tar sample, illustrating the complexity of tars. Publications such as those by Novotny et al. [1981] and Burchill et al. [1983] provided a starting point on coal tar composition to aid the development of later research.

The most extensive published study on coal tar composition to date is research conducted by Brown

et al. [2006]. Comparative analysis of the chemical and physical properties of eleven coal tars was performed. However, the study was limited to the traditional tiered analytical approach, and as such was restricted mainly to the priority PAHs. The differences in composition across the various sites was not discussed and the study focussed on environmental implications rather than source apportionment. Thus, the novel method for chemical fingerprinting of coal tars as developed in this project is of great benefit for environmental forensic investigations at impacted FMGP sites.

## 2.8   Environmental Forensics at Former Gasworks

Although a considerable amount of literature has been published on coal tar composition, very little attention has been given to source allocation of tars. The analysis and interpretation of coal tar data is particularly challenging as these contaminants have been exposed to the subsurface environment for decades, if not centuries, therefore a significant degree of weathering is likely.

The Electric Power and Research Institue (EPRI) has an extensive catalogue of technical reports applying traditional oil spill fingerprinting techniques to coal tar contaminated samples. Within one such report, Mauro [2000] compared the composition of coal tar and crude oil, demonstrating that both contaminant sources contain similar major components, but in varying concentrations. Therefore, the report investigates PAH diagnostic ratios for coal tar source allocation, equivalent to those used within oil spill fingerprinting. Mauro [2000] demonstrated the difference between coal carbonisation tars and CWG tars through the use of a double ratio plot. Figure 2.14 depicts this method of classification of tar samples using alkyl phenanthrene and alkyl dibenzothiophene ratios commonly used in oil-oil and oil-source correlations.
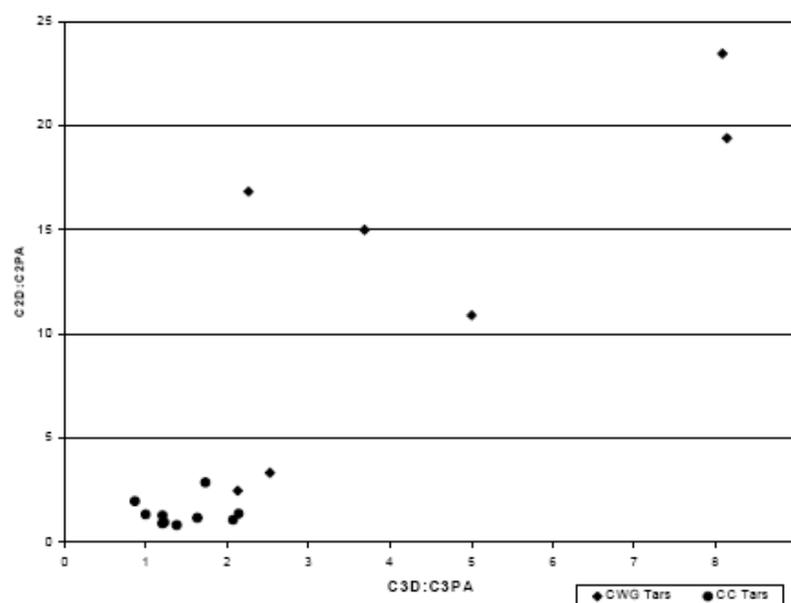
Figure 2.14: Diagnostic ratio plot for coal tar classification, where C2D & C3D are the C2 and C3 alkyl dibenzothiophenes, C2P and C3P are the C2 and C3 alkyl phenanthrenes, CWG = carbureted water gas and CC = coal carbonisation (copied from Mauro, 2000).

However, the study was very limited, with investigation of only two different types of coal tar. Mauro [2000] states that it is only recently, with the redevelopment of FMGP sites, that site owners have become aware of the complexity of the contamination, as there may be multiple coal tar DNAPL plumes located across a single site. Nevertheless, the report does not provide a method for accurate determination of coal tar source and shows only vague trends between tar formed by two different processes.

More recently, Emsbo Mattingly et al. [2006] analysed numerous coal tar types in a GC-MS study. The research focussed on the interpretation of data by chromatographic pattern recognition and diagnostic ratios. The chromatograms in Figure 2.15 demonstrate that there may be significant differences in chemical distribution patterns in coal tar samples produced by different manufacturing methods. However, as Figure 2.15 shows, the chromatograms of three different coal retort tars are extremely similar. The study by Emsbo Mattingly et al. [2006] does not allow source-specific characterisation of coal tars. Tars produced by the similar manufacturing processes are likely to have similar chromatographic patterns, thus accurate differentiation of such sources at a single site would be impossible via this method. The novel ultra-resolution approach to chemical fingerprinting

described in this thesis has the potential to resolve such issues surrounding source apportionment of coal tar.



Pyrogenic Features: (1) naphthalene, (2) methylnaphthalenes, (3) phenanthrene, (4) fluoranthene, (5) pyrene, (6) chrysene, (7) benzo(a)pyrene, and (8) benzo(g,h,i)perylene
Petrogenic Features: UCM is Unresolved Complex Mixture
Quality control compounds (Q)
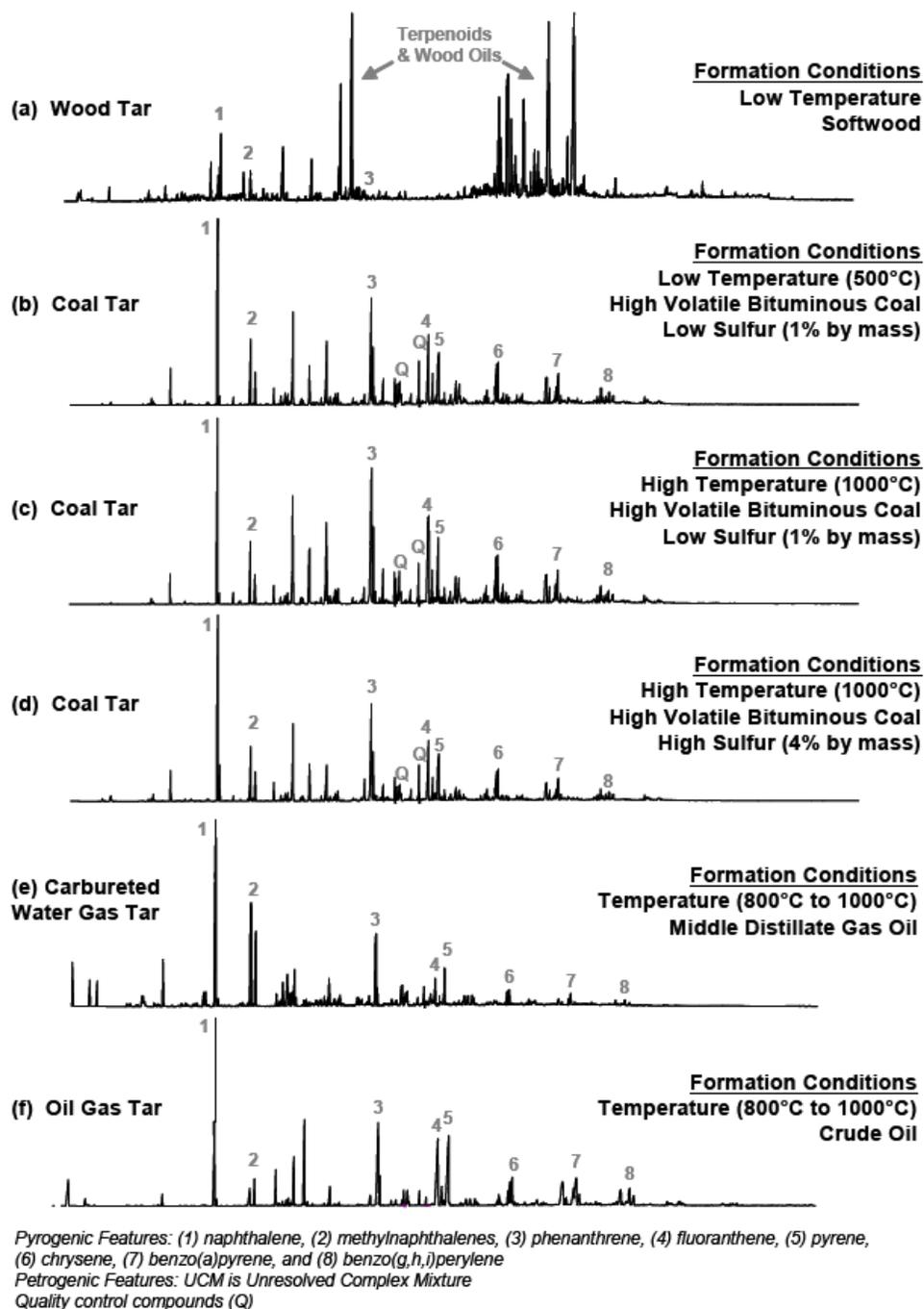
Figure 2.15: GC-MS chromatograms of various tar types showing similarities in PAH distributions (copied from Emsbo Mattingly et al., 2006).

## 2.9   Conclusion

In this Chapter, the historic processes employed at former manufactured gas plants were presented and correlated with known changes in by-product composition, which in turn affects site contamination. The chemical nature of coal and particle size, pyrolysis temperature and catalysis by the hot retort walls may all affect reaction rates and the mechanism of coal tar formation, resulting in contaminated sites of high complexity.

The traditional method of chemical fingerprinting and its application to coal tar was reviewed. It was demonstrated that conventional analytical procedures are not adequate for accurate source allocation of coal tars during environmental forensic investigations. Additionally, the statistical interpretation of results has been limited to simple, ratio plots which are not capable of depicting process-specific differences in coal tar composition.

The aim of this thesis is to develop a comprehensive method of coal tar analysis suitable for environmental forensic applications. In the following Chapter, the principles of the analytical and spectroscopic instruments used for this research are discussed in detail.

# Chapter 3

# Analytical and Spectroscopic Techniques

## 3.1 Introduction

The aim of this project was the development of an accurate system of chemical fingerprinting through utilisation of a suite of advanced analytical techniques. It is for this reason that a chapter has been included to describe the principles of the analytical instrumentation used within the project. The specific instrumentation and parameters used will be discussed in Chapter 4.

## 3.2 Sample Preparation Techniques

### 3.2.1 Accelerated Solvent Extraction

Accelerated solvent extraction (ASE) is an automated method for the batch extraction of solid and semi-solid sample matrices. The technique uses elevated temperatures and pressures to aid the breakage of analyte to sample matrix bonds [Richter et al., 1996]. Elevated temperatures cause decreased solvent viscosity to provide increased penetration of the solvent into the sample matrix.

Elevated pressures are employed to ensure the solvent remains in the liquid phase throughout the extraction [Richter et al., 1996].

Samples are packed within individual, stainless steel extraction cells. Each extraction cell is lined with filter paper to ensure unwanted particulate matter does not collect in the final sample. The sample is then added to the cell and any remaining cell volume packed with an inert solid, such as diatomaceous earth (DE). A portion of activated silica gel can also be included below the sample to provide simultaneous extraction and cleanup.

The packed extraction cell is placed onto an autosampler carousel which transports the cell into a pre-heated oven. After the designated heating time, the cells are pressurised with nitrogen gas (to approximately 10 MPa) and the solvent is injected at the top of the cell. The resulting sample extract is forced from the base of the cell into a collection vial. Extractions may be performed in either dynamic or static extraction mode, or a combination of the two modes. In static extractions, the solvent remains within the cell until the extraction reaches equilibrium, then is rapidly flushed into the collection vial by further solvent and high pressure. In dynamic extraction mode, the fresh solvent continuously flows through the cell at a controlled rate [Richter et al., 1996].

### 3.2.2   Sample Evaporation

Prior to instrumental analyses, environmental extracts may require concentration by evaporation of excess solvent. This was achieved using a parallel work-up station consisting of four main sections; a heated sample rack (with shaking mechanism), a vacuum pump, a condenser and a freezer unit. Multiple samples can be concentrated simultaneously by placement in suitable vials within the sample rack. The heated rack and vacuum pump encourage the evaporation of solvent, which is then collected by a condenser unit. A freezer unit is required to circulate coolant around a small aliquot of the sample held within a tapered section at the base of each sample flask; this prevents the sample from being evaporated to dryness. In addition, the sample rack incorporates a shaking mechanism to ensure a strong vortex in the sample is produced, thereby preventing evaporation retardation. The temperature and pressure are controlled by the operator and are chosen based on the volatility of sample solvent.

## 3.3   Capillary Column Gas Chromatography

### 3.3.1   Theory of Chromatography

The IUPAC (International Union of Pure and Applied Chemistry) definition of chromatography is as follows:

"...a physical method of separation in which the components to be separated are distributed between two phases, one of which is stationary (the stationary phase) while the other (the mobile phase) moves in a definite direction" [Ettre, 1993].

Chromatographic techniques are generally classified according to the type of phase used in the system. For example, gas-liquid chromatography (GLC) employs a gaseous mobile phase known as the carrier gas and a liquid stationary phase. In capillary column gas-liquid chromatography, the stationary phase is a thin layer of liquid coating the walls (usually composed of glass or fused silica) of a long, narrow tube known as a column. The term gas chromatography (GC) is now commonly used to describe all chromatographic systems in which the mobile phase is a gas.

**Retention Time**    The time taken for a component to elute at the detector at the end of the column is deemed the retention time, $t_R$ (Figure 3.1). The retention factor, k, describes the ability of a stationary phase to retain a molecule (Equation 3.3.1) where $t_M$ is the dead time; the time taken for the carrier gas to pass through the length of the column [Scott, 2003]. Theoretical retention factors calculated for specific carrier gases and column dimensions can be used to indicate if the system is leak tight.

$$k = \frac{t_R - t_M}{t_M} \tag{3.3.1}$$

The components of a sample mixture partition between the stationary phase and the carrier gas, based on their solubility in the liquid stationary phase [Scott, 2003]. Molecules with a greater affinity for the stationary phase will take longer to reach the detector than those which preferentially interact with the mobile phase. The distribution coefficient, $K_c$, defines the equilibrium constant between the two phases (Equation 3.3.2).
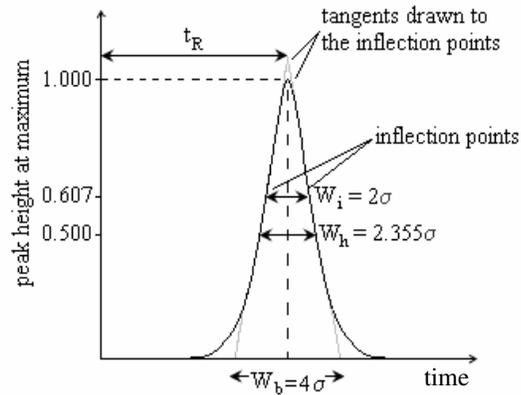
Figure 3.1: Widths of an ideal Gaussian peak as a function of the standard deviation of the peak (copied from Ettre, 1993).

$$K_c = \frac{W_{i(s)}/V_s}{W_{i(M)}/V_M} \qquad (3.3.2)$$

where $W_{i(S)}$ and $W_{i(M)}$ are the amounts of component $i$ in the stationary and mobile phases respectively, while $V_S$ and $V_M$ are the volumes of the stationary and mobile phases, respectively [Ettre, 1993].

However, equilibrium is not established instantaneously along the length of the column, which results in a bell-shaped (Gaussian) peak signal; a typical Gaussian curve is shown in Figure 3.1 [Scott, 2003]. The theoretical plate model, a concept initially developed for distillation processes, was adapted for chromatographic separations to provide a mathematical model explaining the ideal Gaussian peak shape [Scott, 2003]. The model views the column as being divided into a number of theoretical plates (N). Equilibriation of the analytes between the mobile and stationary phase occurs separately from one plate to the next, giving a slight retention of the analyte [McNair and Miller, 2009].

The plate height (H) or height equivalent to a theoretical plate (HETP) can be calculated using the column length (L) and the number of theoretical plates (Equation 3.3.3).

$$HETP = \frac{L}{N} \qquad (3.3.3)$$

Greater separation can be achieved by increasing the number of theoretical plates (N) and by reducing the height of the plates [McNair and Miller, 2009]. The number of theoretical plates can therefore provide a measure of column performance or efficiency [McNair and Miller, 2009]. In an ideal Gaussian curve (Figure 3.1), the peak width at the base line ($W_b$) is described as being four standard deviations ($\sigma$) long and the variance of the peak can be given by the square of the standard deviation ($\sigma^2$) [McNair and Miller, 2009]. The number of theoretical plates is related to the variance in the Gaussian curve by $N = \frac{L^2}{\sigma^2}$ where the standard deviation ($\sigma$) can be represented by Equation 3.3.4.

$$\sigma = \frac{LW}{4t_R} \tag{3.3.4}$$

Therefore, the number of theoretical plates can be calculated by Equation 3.3.5.

$$N = 5.54 \left(\frac{t_R}{W_h}\right)^2 = 16 \left(\frac{t_R}{W_b}\right)^2 \tag{3.3.5}$$

where $W_h$ is the peak width at half height and $W_b$ is the peak width at the base line [McNair and Miller, 2009]. $N$ is unitless, as both the numerator and denominator in Equation 3.3.5 have the same units, typically time or distance units, linked by the velocity of compounds in the column, $L/t_R$ [McNair and Miller, 2009]

**Resolution**    The separation of two eluting bands can simply be described by the separation factor ($\alpha$); the ratio between retention factors of two different compounds (Equation 3.3.6).

$$\alpha = \frac{k_A}{k_B} \tag{3.3.6}$$

The separation factor describes the difference in retention factor of a compound with low retention factor ($k_A$) to that of a more highly retained compound ($k_B$). However, this expression does not account for variations in peak width which would affect the separation.

In chromatographic systems, the degree to which two adjacent peaks separate is deemed the resolution ($R$). The resolution is described by:

$$R = \frac{2d}{W_A + W_B} \qquad (3.3.7)$$

where $d$ is the difference between peak maxima of compound A (least retained) and compound B (most retained) while $W_A$ and $W_B$ describe the peak widths of compounds A and B respectively [McNair and Miller, 2009].

The resolution may also be expressed in terms of the efficiency, retention factor and selectivity factor by incorporating Equations 3.3.1, 3.3.5 and 3.3.6 into Equation 3.3.7 to give:

$$R = \frac{1}{4}\sqrt{N}\left(\frac{\alpha - 1}{\alpha}\right)\left(\frac{k_B}{k_B + 1}\right) \qquad (3.3.8)$$

### 3.3.2   Optimisation of Chromatographic Separation

The chromatographic separation can be optimised by modifying the number of plates or the distribution coefficient of the system. In this section, the parameters which allow modification of these values will be discussed.

**Effect of flow rate and carrier gas**   The number of plates depends on the extent of peak broadening in the chromatogram. Peak broadening occurs because the rate at which the equilibrium in the theoretical plates established is finite.

The Van Deemter equation (Equation 3.3.9) predicts that there is an optimal flow rate at which there is minimum variance per unit column length and, hence, a maximum efficiency [McNair and Miller, 2009].

$$HETP = A + \frac{B}{u} + Cu \qquad (3.3.9)$$

where $u$ is the average linear velocity of the mobile phase.

In capillary GC, the eddy diffusion factor (A) is zero as this parameter accounts for different paths a molecule may take through a packed column; the path is linear thus the Golay equation (Equation 3.3.10) was created [McNair and Miller, 2009].

$$HETP = \frac{B}{u} + C_s u + C_m u \tag{3.3.10}$$

The longitudinal dispersion is represented by $B/u$. The analytes disperse from areas of high to low concentration i.e. from the centre of the column towards the edges. However, when linear velocity ($u$) of the mobile phase is high, there is less available time for diffusion to occur and B has a limited effect on the number of plates [McNair and Miller, 2009].

Finally, $C_s u$ and $C_m u$ represent the resistance to mass transfer in the stationary phase and mobile phase respectively. In other words, if the carrier gas flow is too high, the analyte present in the mobile phase will move ahead of the same analyte within the stationary phase, resulting in broadening of chromatographic peaks. A comparison of Golay curves for three common carrier gases is presented by Figure 3.2. Each curve has an optimal value representing ideal chromatographic conditions. The operator can attempt to achieve optimal conditions by controlling the carrier gas flow in the column. The linear velocity and flow rate (q) are linked by Equation 3.3.11.

$$q = u\pi \left(\frac{d}{2}\right)^2 \tag{3.3.11}$$

where $d$ is the internal diameter of the column.

In capillary column GC, the flow rate is determined by the pressure of the carrier gas at the inlet to the column (Equation 3.3.12).

$$q = \frac{\pi d^4}{128\eta} \left(\frac{P_i^2 - P_o^2}{2P_o}\right) \tag{3.3.12}$$

where $\eta$ is the dynamic viscosity, $P_i$ is the inlet pressure and $P_o$ represents the outlet pressure.

As demonstrated by Figure 3.2, hydrogen and helium give low resistance to mass transfer and high diffusivities making them very popular choices as a GC carrier gas.
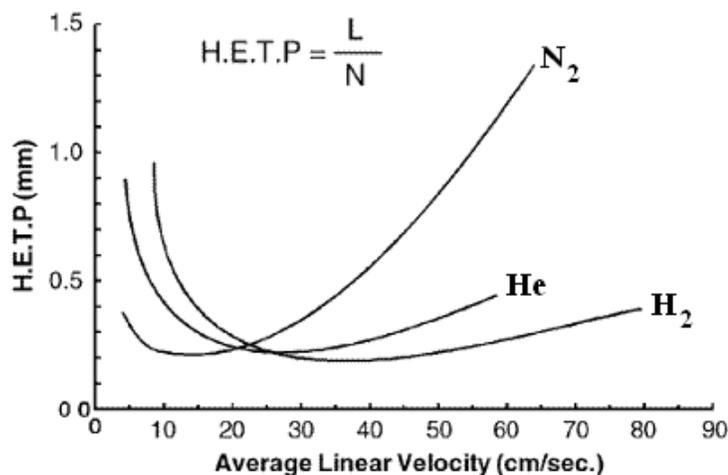
Figure 3.2: Effect of carrier gas on Golay curve (copied from Restek, 2008).

**Column choice**   The selection of an appropriate stationary phase for a particular sample is a crucial step for optimising resolution. Stationary phases can be classed roughly into three groups: non-polar, mid polarity and polar. Columns classifications are based on the polarity and abundance of functional groups within the stationary phase.
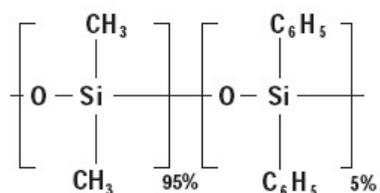


Figure 3.3: Structure of (5%-Phenyl)-methylpolysiloxane stationary phase, as found in DB-5 (or equivalent) capillary columns.

The principle "like dissolves like" enables the operator to predict how a sample will behave within a particular column. For example, a sample containing a high aliphatic (non-polar) hydrocarbon content would be suited to separation by a non-polar stationary phase. A typical non-polar stationary phase, which works by Van der Waals retention of compounds, is presented in Figure 3.3. Non-polar columns are more thermally stable thus it is recommended that the least polar column capable of separation is chosen. Additionally, column parameters such as internal diameter, length

51

and stationary phase film thickness must be optimised to improve resolution. For example, a thick stationary phase film will retain components longer than a thin film.

**Temperature programming**   Further optimisation of separation can be achieved by careful programming of the chromatographic oven temperature. As previously mentioned, it is possible to vary the temperature of the oven throughout the analytical run to encourage elution of analytes; the solubility of analytes in the mobile phase is temperature dependent [Grob and Barry, 2004]. The higher molecular weight (low volatility) analytes have a greater affininty for the stationary phase rather than the mobile phase and thus have longer elution times. Increasing the oven temperature can increase the volatility of the analytes and encourage their elution from the column [Grob and Barry, 2004]. However, co-elution of analytes may occur if the temperature ramp rate is too fast, thus the optimisation of temperature program is an integral part of optimising chromatographic resolution.

### 3.3.3   Gas Chromatographic Instrumentation

The key components of any gas chromatographic system are presented in Figure 3.4. The sample is injected into a heated port, where the components are vapourised and transported onto the column by the carrier gas; consequently the application of GC is limited to analytes which are volatile enough to vapourise within the injection port. The column is held within an oven to allow both isothermal (at a single temperature) and temperature programming (gradual increase in temperature) analyses [Grob and Barry, 2004]. The compounds elute from the column into the detector, where a signal is generated and represented as a peak in a chromatogram.
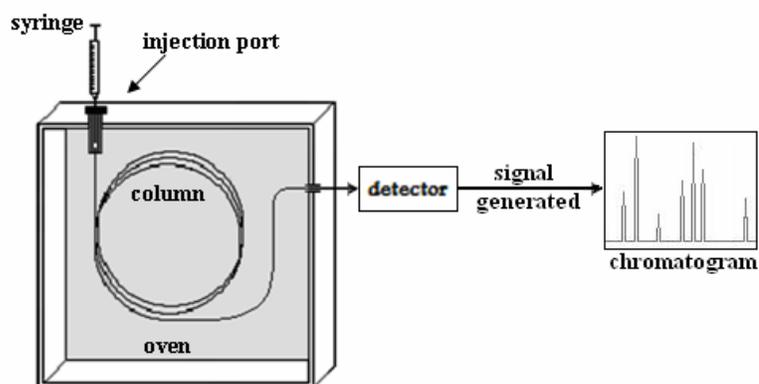
Figure 3.4: Basic components of a gas chromatograph.

**Sample Injection**    All gas chromatographs used for this research were fitted with split/splitless injection ports; meaning the injector can be operated in either split or splitless modes. A schematic of a typical split/splitless injection port is provided in Figure 3.5. The sample enters the glass liner of the port by syringe injection through a rubber septum. The septum purge valve ensures a flow of carrier gas constantly flushes away any volatiles produced by heating of the septum, thus preventing their entry onto the column [McNair and Miller, 2009]. The injection port temperature is controlled by the operator and is generally dependent on the injection technique and sample type.

In split mode, only a portion of the sample is allowed to enter the column, while in splitless mode the entire sample is analysed. The injection port contains a split valve, controlled by flow rate, which can be programmed to open a set time period after injection of the sample [McNair and Miller, 2009]. When the sample is injected into the heated GC port, the molecules begin to vaporise and travel onto the column. When the split valve (or split vent) opens, the injection port is flushed out and prevents further sample from entering the column. Therefore, the longer the split valve remains closed, the more material will be deposited onto the column. The split ratio (e.g. 50:1, 150:1) represents the ratio between the split valve flow and the column flow. The use of split mode allows concentrated samples to be diluted by the carrier gas prior to entering the column. When switching between split and splitless injections the glass liner held within the port must be replaced. The type of liner used for split injections consists of a straight tube ideal for homogenisation of gases prior to injection onto the column. However, in splitless mode, a glass liner with a tapered inner diameter is employed to focus the entire sample onto the column. All GC analyses in this project were performed using split injections.
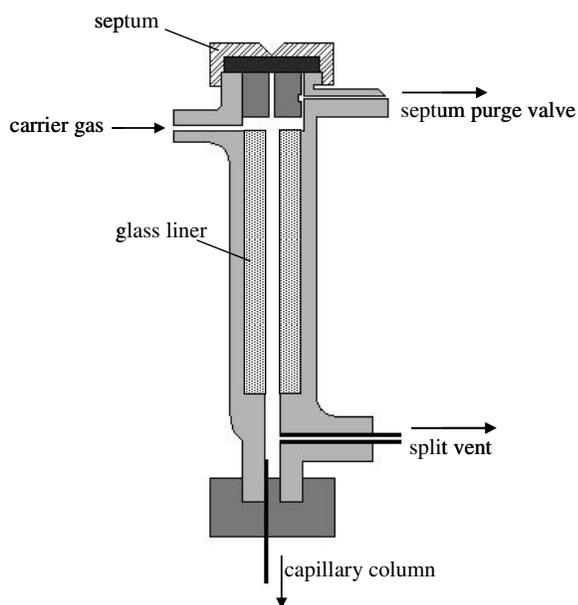
Figure 3.5: Schematic of a split/splitless GC injection port (adapted from McNair and Miller, 2009).

**Liquid Injections**   All injections performed in this study were of liquid samples. Headspace injection of gaseous samples may also be used, however, they will not be discussed in detail here. In liquid injections, a liquid sample is drawn directly into a syringe and injected through the rubber septum of the injection port. The volume injected depends entirely on sample concentration, column capacity and split ratio. Liquid samples require vaporisation within the injection port, hence high temperatures must be used. The injection temperature is chosen based on the solvent and the least volatile compounds present in the sample.

**Non-specific detection**   A variety of detectors are available for GC analyses, the simplest of which are non-specific (or non-selective) detectors. These detectors have a similar response for a broad range of analytes, however, it is not possible to identify individual compounds other than through chromatographic comparison with standards [McNair and Miller, 2009]. An example of non-selective detection is flame ionisation detection (FID). The technique is deemed non-selective as it is able to detect a wide range of organic compounds with only a few exceptions, e.g. non-combustible gases, such as water and carbon dioxide [Scott, 2003]. FID works by combining the eluting constituents (in the gas phase) with a stream of hydrogen and directing this into a chamber

through which air is passed. The hydrogen is ignited in a continuous flame and as the sample constituents pass through it they undergo combustion, producing ions and electrons. A difference in potential of approximately 300 volts is applied to two electrodes situated nearby the flame [Scott, 2003]. As the ions flow between the two electrodes a current is generated and recorded as the chromatographic signal. The intensity of the current produced is directly proportional to the number of ions produced (i.e. the amount of analyte present).

For environmental forensic applications, the level of information obtained by GC-FID is often not adequate for source apportionment. Following the traditional tiered approach to chemical finger-printing, the next step involves mass spectrometry to allow identification of specific compounds (such as biomarkers) within the mixture.

## 3.4   Gas Chromatography Mass Spectrometry

### 3.4.1   Mass Spectrometry

Mass spectrometry (MS) is a spectrometric technique used to separate charged species according to their mass-to-charge (m/z) ratio. The number of ions of each m/z are recognised by a detector and are represented as a mass spectrum which can be interpreted to resolve the chemical structure of unknown samples [Scott, 2003].

All mass spectrometers consist of five main sections: sample inlet, ionisation source, mass analyser, ion detector and a data processing system [de Hoffman and Stroobant, 2007]. A vast number of commercial MS systems exist, however, the only type discussed here will be quadrupole mass spectrometry as this was the only form of MS used in conjunction with gas chromatography in this work. Later sections will describe the principles of other mass spectrometers which were used in conjunction with two-dimensional gas chromatography and for isotopic analyses.

**Electron Ionization**

The principle of quadrupole MS was developed in 1953 by Paul and Steinwegen at Bonn University [de Hoffman and Stroobant, 2007]. Quadrupole mass spectrometers are now commercially available

within a number of different systems and produced by various manufacturers [de Hoffman and Stroobant, 2007]. The first stage in any MS system is the conversion of the sample into gaseous ions. There are a number of ways to ionize a sample, but for the purpose of this research only electron ionisation was used, thus will be discussed in depth.

Electron ionisation (EI) is one of the most common ionisation techniques used in mass spectrometry; a schematic of the system is shown within Figure 3.6. Electrons are produced by passing a charge through a filament. The electrons are accelerated into the ion source chamber by an electric field to create a high energy beam (conventionally set to 70 eV). Analyte molecules are exposed to this beam causing them to lose a valence electron and become a radical cation ($M^{+\cdot}$). Some molecules only undergo this initial loss of an electron and are referred to as the molecular ion, while others have sufficient energy to undergo further fragmentation. A mass spectrum of the possible product ions is produced and can be used to identify unknown compounds. The conventional use of a 70 eV electron beam has allowed a database of mass spectra to be compiled using inputs from many mass spectrometrists [de Hoffman and Stroobant, 2007]. This database by the National Institute of Standards and Technology (NIST) allows users to compare unknown samples to a reference library of spectra.
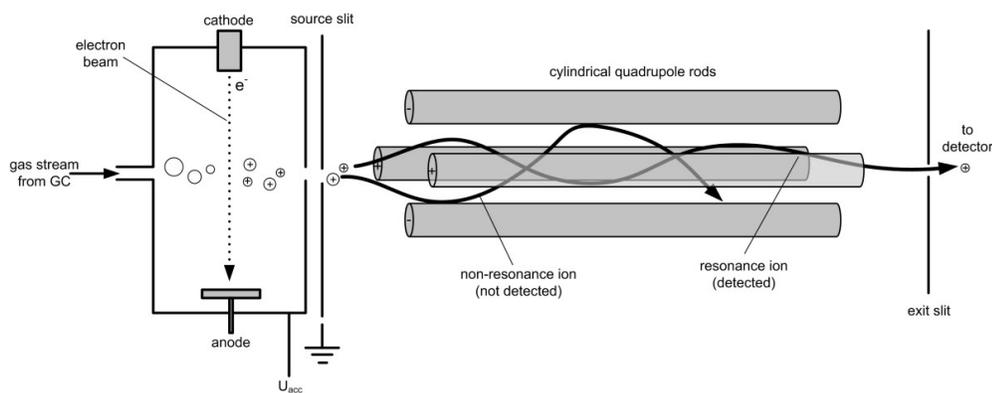


Figure 3.6: Schematic of a mass spectrometer using electron impact ionisation with a quadrupole mass filter (reproduced from Wittman, 2007)

**Quadrupole Mass Spectrometer**

A quadrupole mass analyser is composed of a set of four parallel metal rods, with a direct currect (DC) potential and high radio frequency (RF) signal applied across opposite rods (Figure 3.6). Ions exiting the source are accelerated through the central space between these rods. This causes ions with a particular m/z value to pass straight through to the detector, while other ions have unstable trajectories towards the rods [Scott, 2003].

The four rods are connected in opposite pairs (Figure 3.7) with an RF potential applied between the two pairs. The RF potential alternates to focus the ions in different planes. In other words, a positive ion entering the space between the rods will be attracted to a negative rod. If the potential changes sign before the ion reaches the rod, due to the oscillating electric fields, it will be repelled in the opposite direction. The DC potential and RF signal can be altered to allow ions with different m/z values to reach the detector in turn [Scott, 2003].
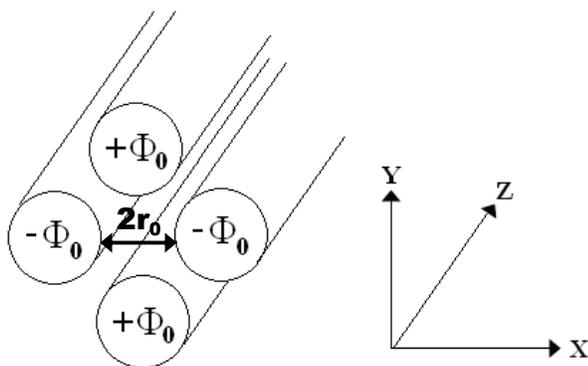


Figure 3.7: Cross section of quadrupole rods illustrating applied potentials (adapted from de Hoffman and Stroobant, 2007)

The potential ($\Phi_0$) applied to the rods can be described by:

$$\Phi_0 = +(U - V \cos \omega t) \text{ and } -\Phi_0 = -(U - V \cos \omega t) \tag{3.4.1}$$

where $\omega$ is the angular frequency, $U$ is the DC potential on the rods, $t$ is the elapsed time and $V$ is the "zero-to-peak" amplitude of the RF voltage [de Hoffman and Stroobant, 2007].

The stability of an ion within the quadrupole can be described with respect to its motion in the x, y and z axes by the following:

$$\frac{d^2x}{dt^2} + \frac{2ze}{mr_0^2}[U - V\cos(\omega t)]x \;\; = \;\; 0 \qquad\qquad (3.4.2)$$

$$\frac{d^2y}{dt^2} - \frac{2ze}{mr_0^2}[U - V\cos(\omega t)]y \;\; = \;\; 0 \qquad\qquad (3.4.3)$$

where $m$ is the mass of the ion, $e$ is the electronic charge ($1.602\text{x}10^{-19}$C), $r_0$ is one half the distance between the rods, $x$ and $y$ represent the distance from the origin in the $x$ and $y$ directions respectively [Leary and Schmidt, 1996].

The equations describing an ion's motion in the x and y axes differ only in sign, while the motion in the z axis is zero. This means that the acceleration in the z axis, from the source to the detector, is not dependent on the potential of the rods [Leary and Schmidt, 1996].

If parameters $a_u$ and $q_u$ are defined such that:

$$a_u = a_x = -a_y = \frac{8zeU}{mr_0^2\omega^2} \text{ and } q_u = q_x = -q_y = \frac{4zeV}{mr_0^2\omega^2} \qquad\qquad (3.4.4)$$

The definitions of $U$ and $V$ can then be given by:

$$U = a_u\frac{m}{z}\frac{\omega^2 r_0^2}{8e} \text{ and } V = q_u\frac{m}{z}\frac{\omega^2 r_0^2}{4e} \qquad\qquad (3.4.5)$$

As previously mentioned, the mass-to-charge ratio ($m/z$) is the measurement used to differentiate between compound peaks by providing characteristic mass spectra showing fragmentation patterns. Equation 3.4.5 clearly shows how the mass-to-charge ratio can be calculated within a quadrupole mass spectrometer. It follows that the trajectory of an ion will remain stable if the values of x and y do not reach $r_0$, in other words, if the ion does not hit the rods [de Hoffman and Stroobant, 2007]. The values of $r_0$ and $\omega$ are assumed to be constant, therefore, for a specific m/z, the value of $a$ is dependent only on $U$ (the amplitude of the DC voltage) and $q$ is dependent only on $V$ (the amplitude of the RF voltage). Due to the symmetry within a quadrupole system, the equations can be simplified into the Mathieu stability diagram (Figure 3.8). The system is set to scan close to the apex at $a_z = 0.24$ and $q_z = 0.706$, so for any given m/z at this point, higher mass ions will
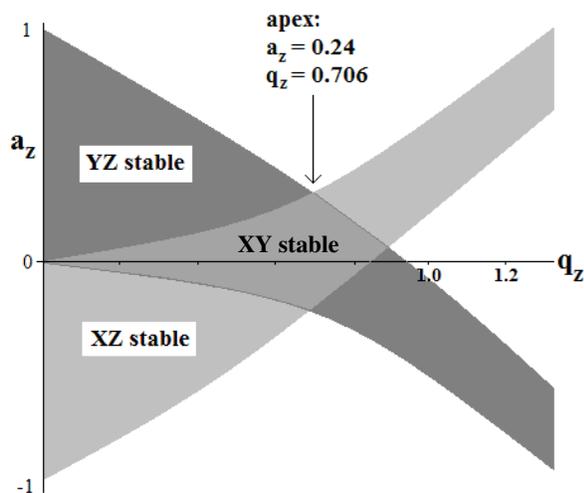
Figure 3.8: Mathieu stability diagram [adapted from Thermo Scientific, 2009].

be unstable at lower $q$ values in one plane and vice versa [Scott, 2003]. In quadrupole MS, $a$ and $q$ (i.e. the DC potential and RF signal) are gradually increased (at a constant ratio) to allow a range of m/z values to be scanned.

## 3.4.2   Coupling of Gas Chromatography with Mass Spectrometry

It is possible to use mass spectrometry as a stand alone technique for chemical identification, however, it is commonly used in conjunction with an online sample preparation method, such as gas chromatography. The hyphenation of gas chromatography with mass spectrometry (GC-MS) allows identification of individual components within a sample mixture. The sample elutes from the GC column directly into the mass spectrometer. A heated transfer line is used to connect the two instruments and to maintain the sample as a gas [de Hoffman and Stroobant, 2007].

Coupling of the two techniques in this manner allows sample components to be identified by both their retention time and mass spectrum. In forensic investigations, two separate lines of evidence are required to corroborate scientific identifications. Gas chromatography coupled with mass spectrometry has the advantage of providing two lines of evidence using a single, analytical technique.

## 3.5 Comprehensive Two-Dimensional Gas Chromatography - Time-of-Flight Mass Spectrometry

### 3.5.1 Instrumentation

Comprehensive two-dimensional gas chromatography (GCxGC) is an advanced analytical technique with an increased separation capacity suitable for the analysis of complex environmental samples [Frysinger et al., 2002]. GCxGC can be regarded as equivalent to planar bed separation techniques, such as thin-layer chromatography (TLC), except that both stages involve gas chromatographic separations [Phillips and Beens, 1999]. The process involves coupling two columns with different stationary phases, to allow separation of a mixture based on two different chemical properties (Figure 3.9). The sample is therefore separated in two dimensions, across a retention plane instead of along a retention line as in conventional gas chromatography. This provides GCxGC with the capacity to resolve an order of magnitude more compounds than traditional gas chromatography [Phillips and Beens, 1999].

As with a conventional GC, the sample is injected into a heated port and forced through the column by a carrier gas. The primary column is generally a long, wide bore (0.25–0.32 mm i.d.), non-polar capillary column, while the second dimension uses a short, narrow bore (0.1–0.2 mm i.d.), polar column; this is deemed normal phase [McGregor et al., 2011a]. However, reversing the column polarity has been shown to provide better group-type separation in certain cases [Van Der Westhuizen et al., 2008]. Reversing the column set in such a way is known as reversed phase (or reversed polarity) GC $\times$ GC. Both columns are held within ovens to allow temperature programming, with the secondary oven generally located within the primary oven (Figure 3.9). The columns are connected by a narrow, glass connector to ensure the entire sample eluting from the primary column will travel onto the secondary column. It is for this reason that GCxGC is deemed a *comprehensive* technique [Frysinger et al., 2002]. The entire sample is subjected to two separations, whereas in *heart-cut* techniques only a selected portion of the sample is trapped and subjected to a second separation [Bertsch, 1999].
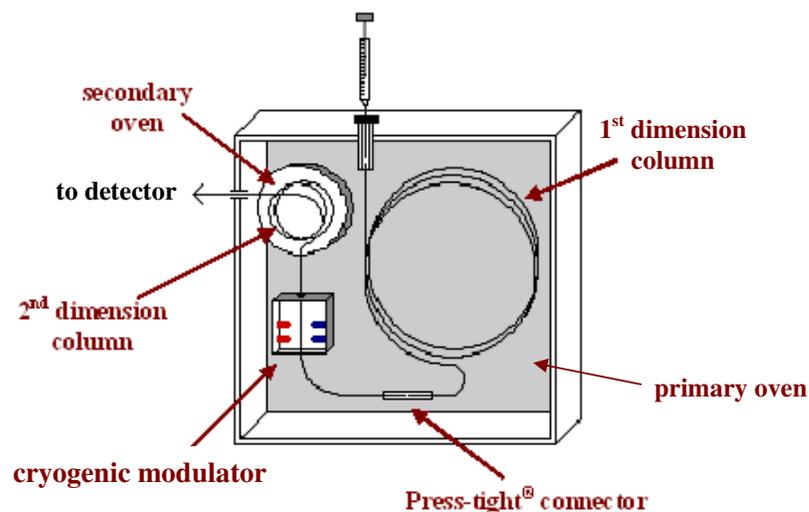
Figure 3.9: Basic components of a GCxGC system [adapted from Leco, 2011].

**Modulation**

One of the most critical parts of the GCxGC system is the modulation device. Every peak eluting from the first column is split into multiple fractions for further separation in the second dimension [Phillips and Beens, 1999]. Separations within the short, narrow secondary columns are very fast, therefore the effluent must be focussed and re-injected in narrow bands to avoid overloading the column. Ineffective modulation results in broad, tailing peaks in the second dimension and can restrict the separation capacity. There are several types of modulation device available with commercial GCxGC instrumentation; namely valve, thermal and cryogenic modulators.

**Valve modulators**   Valve modulators work by precise control of carrier gas flow through a series of valves [Seeley et al., 2000]. The valves periodically direct the sample onto either the second column or to waste, thus causing a pulse of sample into the second dimension. The main drawback of this procedure is the low efficiency, as much of the sample is lost to waste.

**Thermal modulators**   Thermal modulation devices use fluctuations in temperature to retain or desorb analytes from the stationary phase [Phillips and Ledford, 1996]. In this technique, a heating

element rotates around a section of modulator capillary column to provide localised heating and accelerate the effluent onto the secondary column.

**Cryogenic modulators**   The final technique, cryogenic modulation, was the modulation device used within this research. The device uses two closely located pairs of hot and cold nitrogen jets to alternately trap and pulse the first column effluent (Figure 3.10). Initially, the liquid nitrogen-cooled cold jet traps and cryo-focusses the effluent before switching to the hot jet which desorbs the analytes from the stationary phase and push them towards the next pair of jets where the sequence is repeated. This process allows the entire sample to be pulsed onto the secondary column in narrow bands to avoid overloading the secondary column. The columns are most directly connected in series using cryogenic modulation compared to other modulation types [Krupcik et al., 2011]. This is due to the simple column connection employed, with minimal adjustment to pressure and flow at this connection.
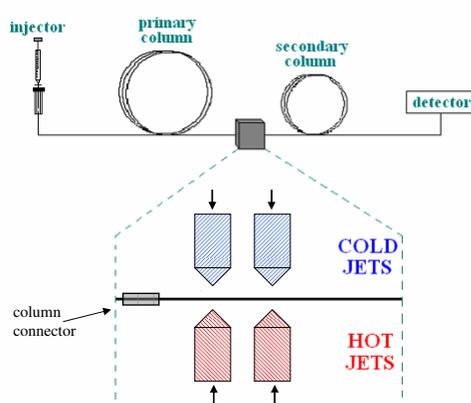


Figure 3.10: Schematic of a cryogenic modulator [adapted from de Koning, 2009].

The output of GCxGC can be represented as a three-dimensional chromatogram by stacking the fast secondary separations side-by-side (Figure 3.11). The results may be viewed/evaluated using this form of chromatogram, however, it is generally easier to compare samples using 2D colour contour plots (Figure 3.12). In a contour plot the x-axis represents the retention time in the primary column, the y-axis represents the retention time in the second column and the colour gradient represents the intensity of the peak, whereas in a 3D surface plot the additional z-axis represents the peak intensity. A contour plot can therefore be thought of as a birds-eye-view of a surface plot, where every spot represents a compound peak separated from the mixture.
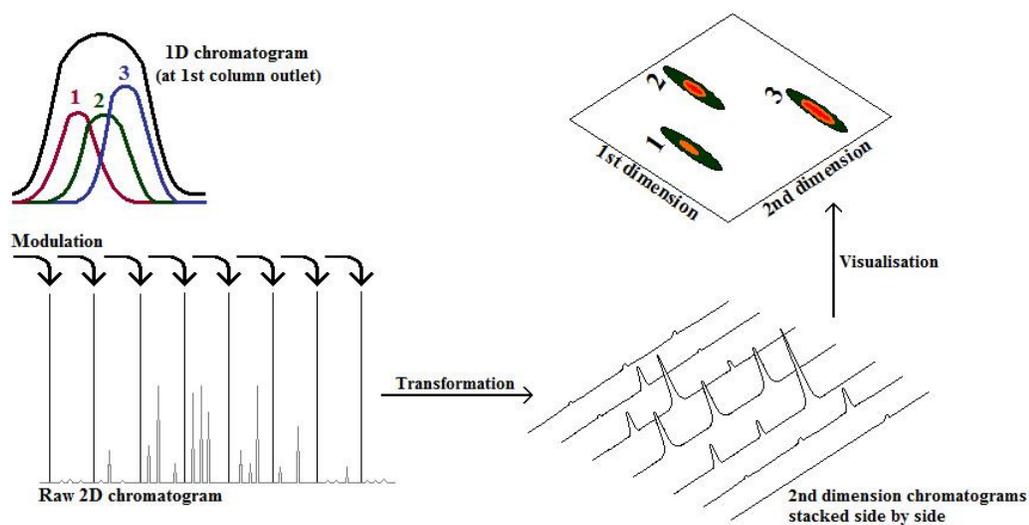
Figure 3.11: Illustration of the production of two-dimensional chromatograms [adapted from Adahchour et al., 2006a].
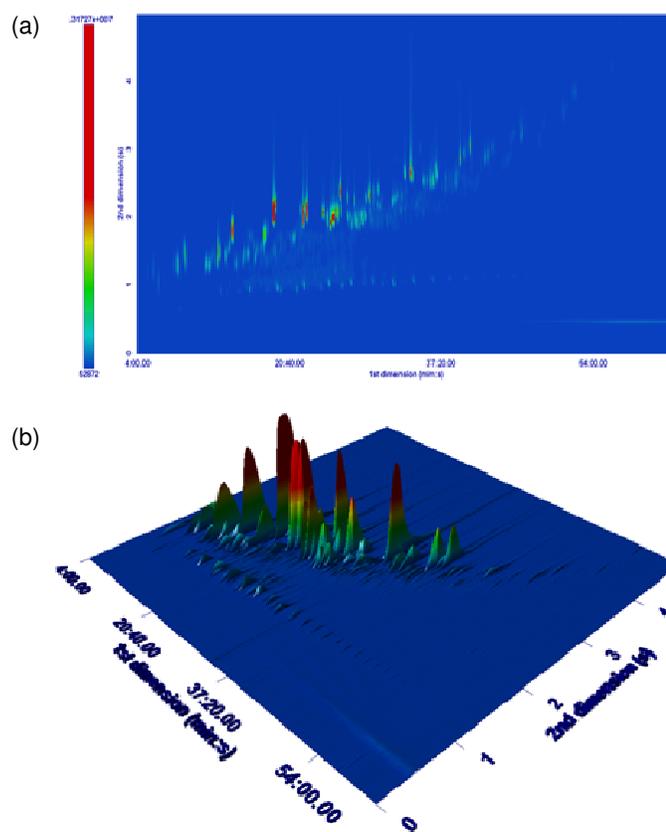
Figure 3.12: Examples of GCxGC chromatographic representations (a) a contour plot and (b) a 3D surface plot [as produced by the analysis of coal tar during the course of this project].

## 3.5.2   Time-of-Flight Mass Spectrometry

**Introduction**

GCxGC can be coupled with non-specific detectors, such as an FID, however, this limits the technique to pattern recognition. Coupling to a mass spectrometer, on the other hand, provides an additional level of information on the sample composition by allowing identification of specific peaks based on chemical structure. Due to the fast separation in the secondary column, a detector with a data acquisition rate of 50-200 Hz is required [Marriott and Shellie, 2002]. Currently, the only type of mass spectrometer capable of meeting these requirements is the time-of-flight mass spectrometer (TOFMS).

Time-of-flight mass spectrometry (TOFMS) determines an ion's mass to charge ratio by their separation in a drift tube after an initial acceleration by an electric field (of known strength). The principle behind the technique is that lighter ions will travel faster than heavier ions, as they require less energy to reach higher velocities, and will therefore reach the detector first. The time taken to reach the detector and the exact length of the drift/flight tube can be used to calculate the mass to charge ratio [de Hoffman and Stroobant, 2007].

**Linear TOFMS**

The example of linear TOFMS will be used to explain the instrumentation involved in time-of-flight mass spectrometry as it is the simplest version of the technique. However, for the purpose of this research, a specific type of TOFMS utilising reflectron technology was employed and will be discussed in detail later. The basic components of a linear TOFMS are illustrated by Figure 3.13; including an ion source, flight tube and a detector.

**Ion Source**   In the ion source, the sample is ionised and the created ions are expelled towards the flight tube by application of an electric field. Within an electric field, ions accelerate in the direction opposite to their polarity (i.e. positive ions will move away from a positive electrode). The negative electrode is composed of a grid, allowing the ions to accelerate through into the flight tube. However, all ions must exit the ion source at the same time to allow correct time/velocity measurement within the flight tube. This is achieved by creating the ions when the positive and negative electrodes are at the same potential. The voltages are then switched on rapidly by a pulse generator to accelerate the ions simultaneously. While these ions are in the process of being extracted, ionization halts to ensure only ions already present in the source can exit through the negative grid. The ions then pass through an accelerating grid which has a much higher voltage potential to accelerate the ions at a velocity suitable for entry into the flight tube [de Hoffman and Stroobant, 2007].

Due to an initial spatial distribution of the ions, the ions closest to the negative grid will leave the ion source faster than those further away, however the ions closer to the positive electrode will be accelerated at a faster velocity. This is known as the kinetic energy distribution, but after a short distance the ions reach the primary focal point and level out [de Hoffman and Stroobant, 2007].
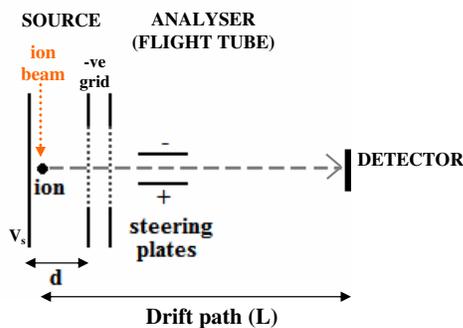
Figure 3.13: Schematic of the basic components of a linear time-of-flight mass spectrometer [adapted from de Hoffman and Stroobant, 2007].

The location of the primary focal point can be adjusted by adjusting grid voltages and spacing between electrodes.

**Flight Tube**    The flight tube is a sealed tube, under vacuum and free of electric fields. When the pressure inside the source is far greater than that required by the flight tube, an additional vacuum point is necessary within the flight tube to remove the influx of gas from the ion source. As the ions drift along the flight tube, propelled by the initial acceleration at the ion source, they begin to separate based on their mass. Once the ions reach the detector they are amplified in turn to produce mass peaks [de Hoffman and Stroobant, 2007]. Additional elements, known as the steering plates, are usually added to ensure that ions which were in motion at the time of creation do not deviate from the axis of the flight tube.

**Detector**    When ions reach the detector plate, they pass through an entry grid signalling the end point of the flight tube. The ions then collide with the detector plate (a micro-channel plate) knocking loose electrons from the surface of the plate. These electrons bounce down the plate disrupting more electrons in a type of domino effect [de Hoffman and Stroobant, 2007]. The electrons are collected by an underlying plate (or anode) which creates a signal in the form of an electrical current, which is ultimately read by a computer.

**Time-of-Flight Theory**

The following section describes how the mass-to-charge ratio can be calculated based on the known constants within the TOF system.    Ions leaving the source will have mass, $m$, and charge, $q = ze$. Ions are accelerated by potential $V_s$, so the electric potential energy ($E_{el}$) is converted to kinetic energy ($E_k$); this relationship is demonstrated by Equation 3.5.1 [de Hoffman and Stroobant, 2007].

$$E_k = \frac{m\nu^2}{2} = qV_s = zeV_s = E_{el} \qquad (3.5.1)$$

The velocity of an ion leaving the source is defined by

$$\nu = (2zeV_s/m)^{\frac{1}{2}} \qquad (3.5.2)$$

It is assumed that the ions travel in a straight line to the detector at a contant velocity, where the time required to reach the detector via the flight tube (path length L) is given by $t = {}^L/v$. The value of $v$ from Equation 3.5.2 can then be substituted in to give Equation 3.5.3, where the mass-to-charge ratio can be calculated by measuring $t^2$, as the terms in brackets are constants [de Hoffman and Stroobant, 2007].

$$t^2 = \frac{m}{z}\left(\frac{L^2}{2eV_s}\right) \qquad (3.5.3)$$

Equation 3.5.3 clearly illustrates that the lower the mass of the ion, the faster it will reach the detector.

**Reflectron TOFMS**

The purpose of a reflectron TOFMS is to improve mass resolution by reducing the kinetic energy dispersion within ions of the same mass-to-charge ratio. As mentioned earlier, the ions located closest to the negative grid will exit the ion source first, but the ions eventually begin to level out at the primary focal point. If the detector was placed here, high resolution could be obtained as the ions will be within a narrow band. However, the distance travelled within the flight tube would be insufficient for mass separation, thus peak overlap would occur [de Hoffman and Stroobant, 2007]. For the remaining length of the flight tube, the faster ions get further away from the slower, heavier ions, leading to widening of the of the cluster and a reduction in resolution.

A reflectron plate has the ability to counteract this effect. In this technique, the ions converge until they meet at the primary focal point and then diverge until they hit a reflector plate (Figure 3.14). The ions are then reflected, converging at a secondary focal point and as they have travelled a further distance in the flight tube by this point the detector can be located here [de Hoffman and Stroobant, 2007]. The goal of the reflectron is to correct the kinetic energy dispersion of ions with the same mass-to-charge ratio. Ions with greater kinetic energy penetrate the reflectron more deeply than those with lower kinetic energy and thus spend more time within it. On entry to the reflectron, the ions pass through a grid with a lower voltage than the source which causes them to decelerate. When the ions drift further into the reflectron they reach an electrode with a higher voltage than the source, causing the ions to stop and accelerate in the opposite direction. Faster ions enter the reflectron further before turning around, while those with a lower kinetic energy will stop and turn around sooner. The faster ions catch up on the way to the detector, thus the reflectron compensates for initial energy distribution.
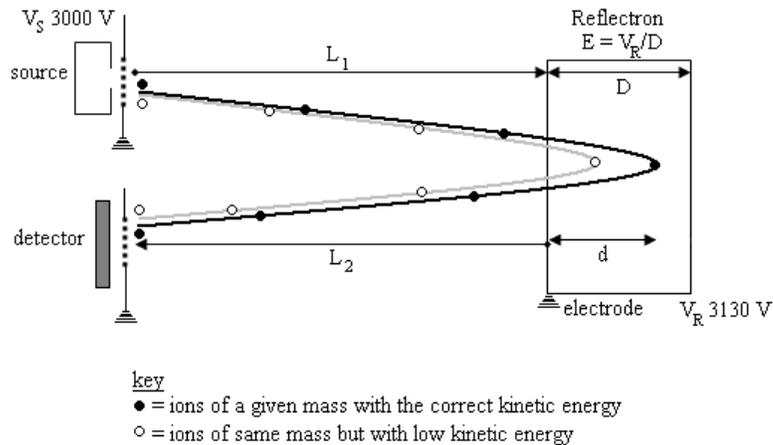


Figure 3.14: Schematic of a reflectron time-of-flight mass spectrometer [adapted from de Hoffman and Stroobant, 2007].

The addition of a reflector plate slightly alters the formula used to calculate the mass-to-charge ratio. Ions of charge $q$ and kinetic energy $E_k$ will enter the reflectron with velocity $v_{ix}$ and penetrate to a depth $d$ (Equation 3.5.4).

$$d = \frac{E_k}{qE} = \frac{qV_S}{qV_R/D} = \frac{V_S D}{V_R} \tag{3.5.4}$$

Equation 3.5.4 illustrates that the penetration depth of an ion in the reflectron is dependent on

the distance between the electrodes within the reflectron and the potential applied on the ions [de Hoffman and Stroobant, 2007]. Due to the higher potential applied within the reflectron than the source, the speed of the ions on the x-axis (i.e. the axis of the flight tube) will eventually be zero and the mean velocity into the reflectron will be equal to $v_{ix}/2$ and the time required will be

$$t_0 = \frac{d}{v_{ix}/2} \tag{3.5.5}$$

The reflectron repels the ions out at the same velocity with which they entered (but in the opposite direction) therefore the distance covered is $2d$ and the time taken is now represented by

$$t_r = 2t_0 = \frac{2d}{v_{ix}/2} = \frac{4d}{v_{ix}} \tag{3.5.6}$$

To calculate the time within the TOFMS, the distance in the flight tube must also be included. Equation 3.5.7 represents the time taken to cover the distance within the flight tube.

$$t = \frac{L_1 + L_2}{v_{ix}} \tag{3.5.7}$$

Therefore, to calculate the total time within the system, $t$ and $t_r$ are combined to give Equation 3.5.8, which once again illustrates how the mass-to-charge ratio can be calculated.

$$t_{total}^2 = \frac{m}{z} \left( \frac{L_1 + L_2 + 4d}{2eV_S} \right)^2 \tag{3.5.8}$$

### 3.5.3   Optimisation of GCxGC separation

To achieve optimal separation in GCxGC it is key to have an understanding of basic chromatographic principles and the instrumentation involved. The parameters for the optimisation of chromatographic separation discussed in Section 3.3.2 still apply, but there are now two separations to consider. This section aims to describe the trends observed between certain instrumental parameters and chromatographic separation by GCxGC. It must be noted that the parameters discussed below correspond to the use of a cryogenic modulator, the only type of modulation technique used within this study.

**Column Sets**

Optimal peak capacity of GCxGC is achieved when the two separations are completely independent; this is known as orthogonal separation (Figure 3.15). Orthogonality is achieved by separation using different chemical mechanisms (i.e. utilising columns with different stationary phases) or by tuning the operating parameters of the secondary column, relative to the progress in the primary column. The choice of stationary phases specific to this research will be discussed in more detail within Chapter 6.
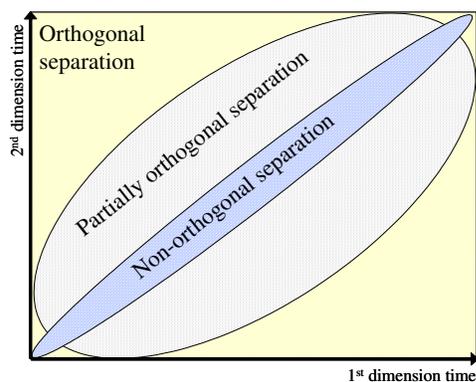


Figure 3.15: Illustration of orthogonality in two-dimensional separations [adapted from de Koning, 2009].

Orthogonality results in one of the main advantages of GCxGC chromatograms; structured ordering of components. Similar chemical classes elute together resulting in a structured chromatogram with a degree of chemical ordering. When a complex mixture is analysed by 1D GC it would be difficult to make assumptions on the chemical structure of eluents based solely on their retention times as they are only separated based on a single chemical property. For example, compounds from many different chemical classes are capable of having similar boiling points, so this alone would not allow classification of different chemical families. However, if these components are further separated based on polarity, as in normal phase GCxGC, classification of chemical families is easier due to the chemical similarities measured by two distinct properties.

Chromatographic ordering allows straight forward interpretation of the results, as chemical classes elute together in bands. For example, the alkylated PAH homologues elute in sequential bands (based on alkylation level) after the parent PAH. This type of structure allows characteristic pat-

terns to emerge and experienced analysts can quickly identify the main chemical classes within a complex mixture.

## Modulation Time

The modulation time is the time taken for the entire cryogenic modulation cycle to be completed. In other words, if the modulation time is set to six seconds this implies that a new portion of the first column eluent will be injected onto the second column every six seconds. The operator has the ability to increase the modulation time to increase the available time of each portion within the second dimension.

In cryogenic modulation, the hot and cold pulse times are directly linked to the modulation time. The modulation time is the total time taken for two cycles of the cold pulse and two cycles of the hot pulse, since the eluents are trapped by two pairs of jets before entering the secondary column. Therefore, if the modulation period is set to six seconds, the time could be split equally between the jets (1.5 seconds each), or the hot jets could be programmed to remain on for a longer time of 2 seconds, meaning that the cold jet time must be set to 1 second to allow a total modulation time of six seconds to be achieved.

The hot:cold jet ratio has a profound effect on the second dimension separation. Compounds which interact strongly with the stationary phase in the second dimension will take longer to pass through the column and may not reach the detector before the next modulation cycle begins; this results in a phenomenon known as *wrap-around* [Mondello et al., 2008]. Wrap-around interferes with the ordered chromatogram as the peaks appear to elute early in the secondary column when they overlap with the next modulation cycle. Increasing the modulation time may prevent wrap-around peaks, however, more often the hot:cold pulse ratio (or the secondary oven/modulator temperature offset) must be increased to encourage the analytes onto the second column.

**Modulation Temperature**   The modulator assembly is maintained at a higher temperature relative to the primary oven to ensure the undisturbed flow of components from the first column onto the second column. A default setting of 15 ℃ temperature offset (relative to the primary oven) is recommended by the instrument manufacturer [Leco, 2008], however, the temperature may have to be increased to increase the speed of elution for strongly retained compounds. Increasing

the modulation temperature offset can prevent wrap-around of peaks in the second dimension. Additionally, the secondary oven temperature offset can also be increased to increase the speed of elution in the second dimension.

## 3.6   Gas Chromatography Isotope Ratio Mass Spectrometry

Isotope ratio mass spectrometry (IRMS) is a technique used to measure the ratio of specific stable isotopes within a sample. Elements of interest are those which have two naturally occurring stable isotopes such as carbon, hydrogen, oxygen, sulfur and chlorine; however, carbon is the most widely used and was the only isotope investigated in this study. A number of review papers have been published detailing the principles and applications of IRMS [Meier Augenstein, 1999, Nic Daeid et al., 2010, Sessions, 2006]. The technique has previously been applied within a number of scientific disciplines including criminal forensics [Benson et al., 2006], geochemistry [Philp, 2007] and environmental chemistry [Slater, 2003].

### 3.6.1   Isotope Ratio Mass Spectrometry

Isotopic ratios are calculated by measuring the ratio of heavy to light isotopes; for example, $^{13}C/^{12}C$ represents the ratio (R) of stable carbon isotopes while $^{2}H/^{1}H$ represents that of hydrogen. The values are always expressed in delta notation ($\delta$) in units of per mille (‰) by measuring the abundance of each isotope within a sample relative to a reference standard of known isotopic value (Equation 3.6.1).

$$\delta(‰) = \left( \frac{R_{sample} - R_{standard}}{R_{standard}} \right) \times 1000 \tag{3.6.1}$$

The variation of carbon isotopes in nature is very low. To show variation within isotope ratios, precision to the third or more significant number is required; not achievable by common mass spectrometers previously discussed within this chapter. Isotope ratio mass spectrometers were developed to measure isotope ratios to the fourth or more significant number [Hoefs, 1980].

Unlike most mass spectrometers, an IRMS is not able to scan a range of m/z values, it is restricted to only those m/z values applicable for stable isotope calculations. This means it has unique collectors for only these specific values and can therefore achieve greater precision and sensitivity [Meier Augenstein, 1999]. The molecules entering the IRMS are not directly analysed, instead they are converted into small gases; $H_2$ and $CO_2$ for hydrogen and carbon isotope analyses respectively (Figure 3.16). This results in a reduced number of collectors required by the IRMS. In carbon analysis mode, three separate collectors are required for the ions of m/z 44, 45 and 46 corresponding to $^{12}C^{16}O_2^+$, $^{13}C^{16}O_2^+$ and $^{12}C^{18}O^{16}O^+$ respectively [Hoefs, 1980]. Commercial IRMS systems are capable of analysing several different isotopes by incorporating multiple sets of collectors (e.g. carbon, hydrogen, nitrogen and oxygen analyses by a single instrument).

To convert the sample into simple gases the system can use two different methods: dual inlet IRMS (DI-IRMS) or continuous flow IRMS (CF-IRMS). In dual inlet systems the gases are prepared offline before admission to the IRMS and includes multiple, time-consuming steps that have a high margin for error. Continuous flow IRMS was the technique use within this study and provides an online method of sample preparation [Benson et al., 2006]. This method uses a helium carrier to force the analyte gases into the IRMS after their production by an online combustion/pyrolysis unit.

Once analyte gases have been formed they are forced into a mass spectrometer. The mass spectrometers used for isotopic analysis consist of three main components: an ion source, a mass analyser and an ion collection assembly [Benson et al., 2006]. As with previously described MS instruments, the gaseous sample enters the ion source where impact with a high energy electron beam occurs, creating positive ions. However, the extraction voltage in this case is optimised for the formation of molecular ions of only the specific gases being measured, instead of remaining at a constant values of 70 eV.

The positive ions are then accelerated out of the ion source and into a flight tube passing between the poles of an electromagnet, which separates the ions based on mass-to-charge ratio. The ions are deflected into circular paths with radii proportional to the square root of their mass-to-charge ratio. Thus, each path corresponds to a separate m/z value [Hoefs, 1980].

The ion beam enters the flight tube with a velocity ($\overrightarrow{\nu}$) due to acceleration out of the ion source

by an electric field. The size ($\nu$) of the velocity is defined by:

$$\nu = \sqrt{\frac{z}{m}2V}$$ (3.6.2)

where V is the potential difference of the source [Hoefs, 1980].

The Lorentzian force law (Equation 3.6.3) can then be applied as the ion beam is passed through a magnetic field perpendicular to the velocity vector and of size B. The Lorentzian force direction is perpendicular to both the velocity and the magnetic field [Hoefs, 1980].

$$\overrightarrow{K} = ze(\overrightarrow{\nu} \times \overrightarrow{B})$$ (3.6.3)

The size of the Lorentzian force is described by:

$$K = ze\nu B sin(\overrightarrow{\nu}, \overrightarrow{B}) = ze\nu B$$ (3.6.4)

This results in the required circular path, as B is constant along the tube, and as the Lorentzian force must be equal to the centrifugal force, the radius (r) of the deflected path can be described using Equation 3.6.5.

$$\frac{m\nu^2}{r} = ze\nu B$$ (3.6.5)

Consequently, the ions with different m/z will be separated into different circular paths within the flight tube and thus will have differing collection points at the detector. Substitution of Equation 3.6.2 into Equation 3.6.5 derives the working equation used to describe an isotope ratio mass analyser (Equation 3.6.6).

$$\frac{m}{z} = \frac{r^2 B^2 e}{2V}$$ (3.6.6)

The ions are collected by a selection of Faraday cup collectors at the end of the flight tube in such a position as to allow simultaneous stimulation by all ion paths (Figure 3.16). The Faraday cups convert the impact of ions into an electrical current which can then be amplified, digitised and transferred to a computer. The computer software then calculates the isotopic ratio by integrating the peak area of each isotopomer. The number of different Faraday cups available depends on the specifications of the instrument in use. Modern instruments are capable of monitoring the abundance of carbon, hydrogen, oxygen and nitrogen isotopes within a single IRMS instrument.
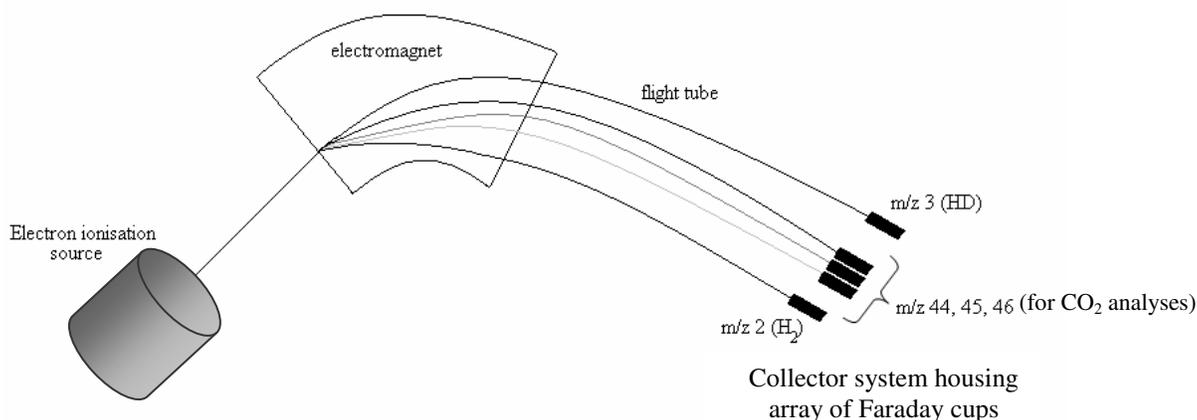
Figure 3.16: Schematic representation of an isotope ratio mass spectrometer [adapted from Schmidt et al., 2004].

## 3.6.2   Coupling to Gas Chromatography

Isotope ratio mass spectrometry is generally used in one of two forms; either bulk stable isotope analysis (BSIA) or compound specific isotope analysis (CSIA). BSIA is limited to averaging the isotopic ratio of all molecules within a sample while CSIA is capable of calculating isotope ratios for single compound within a sample [Benson et al., 2006]. The technique works by coupling a gas chromatograph to an online combustion unit and an IRMS in a process known as GC-combustion interfaced-isotope ratio mass spectrometry or GC-C-IRMS [Meier Augenstein, 1999]. The first systems for online CSIA of carbon were developed in the late 1970s, however, it was not until the 1990s that commercial GC-C-IRMS instruments became available [Meier Augenstein, 1999]. As mentioned earlier, commercial systems are now also available to measure carbon, nitrogen, oxygen and hydrogen isotopes using a single instrument.

The following section focusses on the development of the online combustion units used to continuously and quantitatively convert the injected sample into simple gases for carbon isotope analyses.

**Combustion for Carbon Isotope Analysis**

For carbon isotope analyses, the effluent is forced into a reactor (held within a furnace) for combustion and oxidation of the sample components. A compromise exists between achieving complete combustion and enhancing reactor lifetime; high temperatures provide high combustion capability,

however, they also lead to rapid degeneration of the oxidiser. The most common oxidants in use are copper oxide (CuO) and nickel oxide (NiO). CuO favours low temperatures ($<800$ °C) while NiO has optimal capability at higher temperatures [Sessions, 2006]. Thermo Scientific, manufacturers of the instrument used in this study, developed a solution by combining the use of CuO and NiO within a single reactor, alongside a platinum catalyst, for use with operating temperatures in excess of 900 °C. The reactors must be re-oxidised periodically (after approximately every 50 samples) by flushing with high concentrations of oxygen to retain optimal performance levels. When in full working order, the reactors are expected to quantitatively oxidise all molecules in the sample, thus every carbon atom in the molecule should be converted into a molecule of carbon dioxide. Water molecules are also formed during the combustion process and must be removed by a water-permeable Nafion®membrane to prevent protonation of carbon dioxide molecules.

**Isotopic Calibration**    A reference gas is generally used to allow calibration of a sample against a standard of known isotopic composition. The reference gas ($CO_2$ for carbon analysis) is either introduced into the carrier gas stream or pulsed directly into the ion source. A reference gas is one in which the $\delta$ value has been previously determined against an international standard. The reference gas should be analysed several times within each analytical run to compensate for variations in the mass spectrometer [Hunkeler et al., 2008].

**Chromatographic resolution**    A major challenge within CSIA is the optimisation of chromatographic separation. CSIA relies on baseline resolution of peaks to ensure that the $\delta$ value obtained represents a single compound. Coelution of chromatographic peaks causes inaccuracy of the results, however, there are certain applications of CSIA where this simply cannot be avoided. The chromatographic isotope effect further strengthens the requirement of baseline separation. In carbon CSIA, for example, the heavier m/z 45 ion elutes a fraction of a second earlier than the lighter m/z 44 ion due to differences in chromatographic retention [Meier Augenstein, 1999]. This results in an S-shaped trace of the isotope ratio (Figure 3.17). Partial peak integration can compromise the quality of isotopic data, therefore, integration parameters must be strictly monitored to ensure that both isotopic peaks are included in the detection window. Furthermore, the integration must exclude traces from closely eluting peaks to guarantee high-precision CSIA.
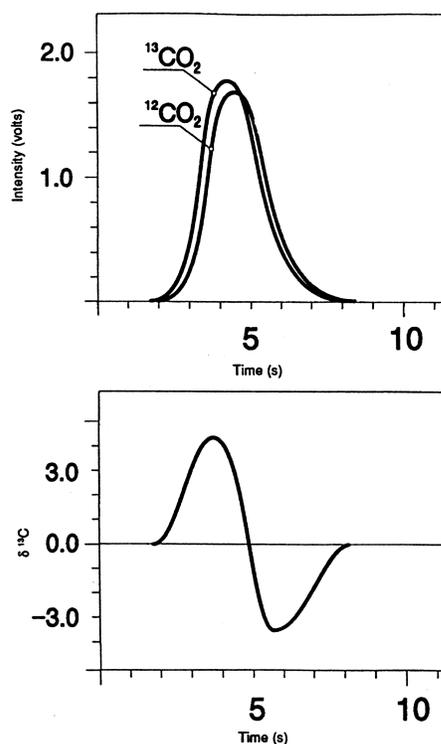
Figure 3.17: Illustration of the time displacement between $^{13}CO_2$ and $^{12}CO_2$ and the resulting S-shaped trace of the isotope ratio [copied from Meier Augenstein, 1999].

### 3.6.3   Quality Control for CSIA

High quality chromatographic separation is exceptionally important in CSIA, however, there are a variety of other quality control procedures which are recommended to ensure that precise and accurate isotope values are obtained. Generally, triplicate injections of each sample are performed, with isotopic results expressed as an average, alongside the standard deviation. Standard lab procedures such as analysis of standard mixtures, blanks and procedural blanks should also be considered [Hunkeler et al., 2008].

**Linearity and Stability**

It is good practice to perform instrument stability checks at the beginning of each analytical sequence. This can be achieved by zero-enrichment tests and linearity checks. A zero-enrichment test analyses a series of reference gas peaks to ensure the standard deviation falls within an acceptable

level. As with all scientific instrumentation, GC-C-IRMS technology is constantly evolving to provide higher quality outputs. The instrumental error involved in CSIA measurements depends on the instrument manufacturer, however, precision of standard deviation less than 0.3 ‰ for $\delta^{13}C$ is generally used by authors [Brenna et al., 1997].

Furthermore, in carbon mode, linearity tests can be performed to verify that non-linearity effects within the ion source are at a minimal level. The term non-linearity effect refers to the ion-molecule reactions which may occur in the ion source and interfere with isotope ratio measurement [ThermoScientific, 2007]. For example, excited $H_2O^{+\cdot}$ ions may react with carbon dioxide molecules as follows:

$$H_2O^{+\cdot} + {}^{12}C^{16}O_2 \longrightarrow {}^1H^{12}C^{16}O_2^+ + OH^{\cdot} \tag{3.6.7}$$
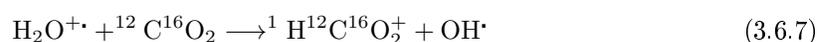
In the mass analyser, the protonated species, ${}^1H^{12}C^{16}O_2^+$, with m/z 45 will interfere with the ${}^{13}C^{16}O_2^+$ ion (also m/z 45) resulting in incorrect isotopic measurements [ThermoScientific, 2007]. Tuning of the ion source conditions can minimise this effect, but careful monitoring by a linearity test is recommended. The linearity test measures a series of reference gas peaks at different intensities and the resulting standard deviation should be less than 0.02 ‰/nA [ThermoScientific, 2008].

### Limits of detection

The peak signal obtained by GC-C-IRMS is directly proportional to the number of atoms of that particular element in the sample. If complete combustion of the sample occurs, each atom of the investigated element should form a molecule of simple gas. For example, in carbon mode each carbon atom in the sample is converted to a molecule of carbon dioxide. Therefore, a molecule of methanol will be converted into a single molecule of $CO_2$, while a molecule of naphthalene ($C_{10}H_8$) should provide 10 molecules of $CO_2$.

As a result, the limit of detection for GC-C-IRMS systems is expressed in moles of analyte element. Low signal-to-noise (S/N) ratio can cause difficulties when defining the peak start and stop points. In carbon analysis, the minimum sample size capable of yielding precisions of SD($\delta^{13}C$) < 0.3 ‰ is approximately one nmole carbon [Brenna et al., 1997].

## 3.7    Quantification

Quantification of an analyte can be performed by gas chromatography. The intensity of the detector response is proportional to the amount of compound detected and thus, the amount of that compound in the injected mixture. Within a range of analyte concentrations, the detector response and analyte concentration can be linearly related. The peak area is generally used as a measure of detector response, though peak height may also be used [McNair and Miller, 2009].

### 3.7.1    Instrumental Calibration

Calibration curves can be produced by injecting a standard solution of target analytes into the GC over a range of concentrations and recording the detector response. The values can be plotted to obtain a linear equation for each analyte, though if the range of concentrations is broad, the resulting curve may be quadratic. Detector response varies for different compounds, therefore, calibration curves are generally produced for each analyte individually.

A known concentration of a compound not naturally found in the sample, referred to as internal standard ($C_{is}$), is added to each calibration standard to ensure that errors in injection are accounted for by normalisation of the peak area of the analyte ($A_s$) against the peak area of the internal standard ($A_{is}$) [USEPA, 1997b].

**Linear Regression**

When two correlated variables, X and Y, are plotted against one another in a graph, a linear equation of the form Y=mX+c can be found. In the case of calibration curves, X represents the analyte concentration while Y represents the analyte peak area. The slope of the line is given by the coefficient, m, and the y-axis intercept is given by c. The values of m and c can be deduced using Equations 3.7.1 and 3.7.2.

$$m = \frac{n \sum (xy) - \sum x \sum y}{n \sum x^2 - (\sum x)^2} \tag{3.7.1}$$

$$c = \frac{\sum y - a \sum x}{n} \tag{3.7.2}$$

where $x$ and $y$ represent measured values of X and Y respectively and $n$ represents the number of measurements. Once a calibration curve has been produced, unknown samples can be quantified by rearranging the linear equation to solve for X (concentration) when Y (peak area) is known.

The level of correlation (or measure of linearity) between the two variables is given by Pearson's correlation coefficient (Equation 3.7.3).

$$r = \frac{\sum (x - \bar{x})(y - \bar{y})}{\sqrt{\sum (x - \bar{x})^2 \sum (y - \bar{y})^2}} \tag{3.7.3}$$

Values of $r$ lie within the range of -1 and +1, with values close to each limit (i.e. -1 and +1) representing strong correlation between the variables. Most statistical software and spreadsheet packages represent an R-squared value during linear regression, which is the squared value of the Pearson's correlation coefficient, and it is this value which will be displayed on calibration curves in this study.

**Quadratic Regression**

In the case of a quadratic calibration curve, the obtained equation is of the form $Y = aX^2 + bX + c$ with fixed constants a, b and c. The concentration of analyte within an unknown sample can be calculated using the constants a, b and c (obtained from the equation of the curve) and the quadratic formula given in Equation 3.7.4.

$$X = \frac{-b + \sqrt{b^2 + 4a(c - Y)}}{2a} \tag{3.7.4}$$

## 3.8   Summary

This chapter has presented the background theory to the numerous gas chromatographic techniques used within this project. The next chapter describes the specific instruments and software employed as well as an overview of the coal tar samples and sites analysed in this study.

# Chapter 4

# Materials and Methods

## 4.1 Samples

Twenty-five coal tar samples (labelled 1-23; with sub-samples 20t1, 20t3 and 20t5 representing sample 20) were provided from fifteen FMGP sites (labelled S1-S15) across the United Kingdom. All samples were provided by Parsons Brinckerhoff (in asssociation with National Grid). Sites were chosen based on the presence of tar (or tar-rich soils) and the existence of historical site data. The approximate location of each site is shown in Figure 4.1 and the basic site details are summarised in Table 4.1. All samples were stored at 4 °C prior to extraction.

### 4.1.1 Historical Site Information

For confidentiality reasons, the sampling locations are not referred to by name; the fifteen FMGP sites sampled in this study are instead labelled S1-S15. The major manufacturing processes over the years of operation for each site are summarised in Table 7.1.

Thorough site histories of each FMGP site in this study exist at the University of Strathclyde (as well as Parsons Brinckerhoff) however, these details are commercial-in-confidence. Nevertheless, it is outwith the scope of this thesis to provide comprehensive site investigations for each FMGP, the focus of this study was to develop a large library of coal tars covering as many historical
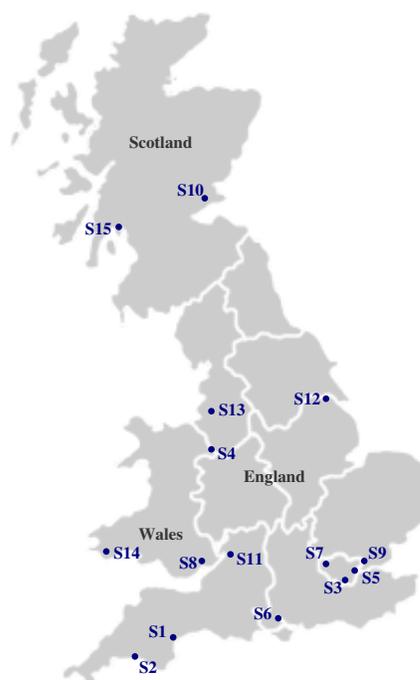
Figure 4.1: Approximate locations of the fifteen former manufactured gas plants sampled in this study (labelled as S1-S15).

manufacturing processes as possible to establish trends and connections for environmental forensic purposes. Thus, the following paragraphs provide general site descriptions, including any relevant details concerning site manufacturing processes.

**Site S1**     Coal tar samples 1-6 were all obtained from an FMGP at site S1 which operated between 1836 until 1971. Initially, the site operated simple horizontal retorts, but the gasworks plan was redeveloped in 1878 after a review by a gas engineer stated it was unsuitable for operation. After redevelopment, the works were expanded in 1912 to include five continuous vertical retort beds. In 1914, annual gas production had reached 434 million cubic feet (cf). This was extended in the 1920s by addition of a further five CVR beds.

In 1942, the site was affected by a World War II air raid, which destroyed a large oil tank containing heavy oil and an ammoniacal liquor plant. From the early 1930s until 1952 the plant remained at capacity, with various improvements to increase efficiency, including the addition of two water gas plants. The site was further expanded in 1952 by construction of a purification plant, additional vertical retorts and the development of storage areas for primary flash distillate (a light petroleum

distillate). It follows that circa 1959 the site began production of gas from oil (mainly primary flash distillate).

The sampling locations are highlighted in a rough site map presented in Figure 4.2. Samples 1 and 2 were taken from boreholes downstream of former gasholders on the site, while samples 3-6 were obtained from an area near the site boundary. Coal tars 3, 4 and 6 were sampled from boreholes within a former tar well, with tars 3 and 4 taken from the same borehole with a one year lag period and coal tar no. 5 sampled from a borehole slightly down gradient of the tar well structure.

The FMGP at site S1 employed a variety of manufacturing processes, encompassing almost all possible production methods. As well as numerous retort and CWG (coal and oil) processes, micro-simplex oil reforming and tar distillation were also performed. The samples from within/nearby the tar tank (samples 3-6) are almost certain to have been associated with the vertical retort house. Samples 1 and 2 were taken from an area downstream of the tank, near the boundary of a suspected coal tar DNAPL plume, so are likely to be of vertical retort origin, but in theory may have been influenced by any of the processes used on site throughout the years of operation.

**Site S2**  Tar sample 7 was discovered at the base of an underground gasholder tank on site S2 during a remediation project in 2009. Little historical information is available on the site; it is believed to have been a simple, country gasworks using low temperature horizontal retorts. However, the gasworks did not close until 1969 and was the last coal gasworks in the area, meaning it is extremely likely that the manufacturing process was upgraded to high temperature horizontal retorts to improve the efficiency of gas production.

**Site S3**   As with the previous sample, coal tar no. 8 was also acquired from a former tar tank during a remediation project. Site S3 was constructed in 1856 and by 1971 it was used solely as a gasholder station. Site plans dated 1932 and 1950 indicate that horizontal retorts and CWG were the major manufacturing processes on site. Coal tar sample 8 was taken from a former 250,000 gallon tank associated with the horizontal retort house.

**Site S4**  Site S4 represents an extremely complex FMGP which operated a number of different manufacturing processes, including various retorts and a CWG plant, until closure in 1953. The
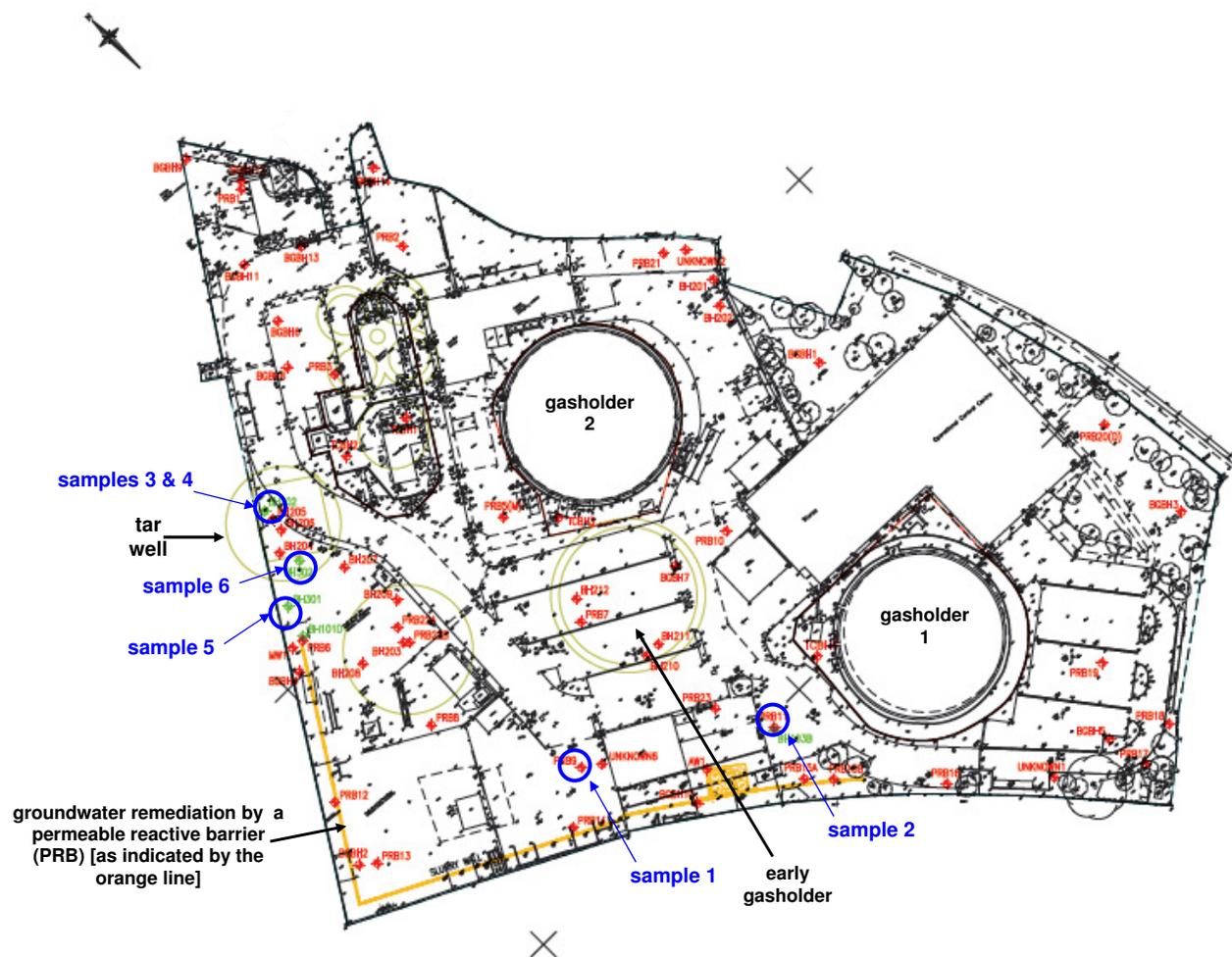
Figure 4.2: Approximate sampling locations of coal tars 1-6 from site S1

CWG plant was commissioned in 1902, while vertical retorts were constructed later in 1913, with additional vertical retorts added in 1931.

Coal tar sample 9 was taken from inside a former tar well during a remediation project. The manufacturing process in use at the time of closure was vertical retort carbonisation, thus it is probable that the tar located in the well was from this process. Nevertheless, the horizontal retorts and CWG plant also located at site S4 may have contributed to the tar.

**Site S5** Coal tar sample 10 was taken from a tar well during a project to remediate the entire former gasworks at site S5. The FMGP at site S5 dates back to 1849, with the major manufacturing process being horizontal retort carbonisation. In 1900, the annual gas production was 253 million cf, which expanded to 525 million cf by 1949.

A site plan of the FMGP (circa 1912) depicts a CWG structure, however, it is not closely located to the tar well sampled in this study. The tar well was located nearby a large retort house and coal storage area, meaning the likely source was a horizontal retort process. The horizontal retort types are likely to have been updated as technology advanced, thus tar sample 10 is likely to be that of a high temperature horizontal retort.

**Site S6** Very little is known of this site other than the fact it was a former creosote works associated with a timber yard. The surrounding area has a long history of industrial processes dating back to the late nineteenth century, including a naphthalene oxidation plant, tar distillery, a bitumen roadstone coating plant and a chemical fertiliser plant. Coal tar sample 11 is likely to be a distilled fraction of coal tar e.g. creosote oil.

**Site S7** Coal tar sample 12 was acquired from a complex former gasworks which opened in 1854 at site S7. Initially, simple (furnace-heated) horizontal retorts were used for gas production, however, by 1957 vertical coal retorts and CWG processes (using coal and oil feedstocks) were also employed. Furthermore, a SEGAS plant was located on site for the conversion of oil to gas as well as a number of different tar tanks associated with tar distillation.

The coal tar sample was taken from a borehole (at a depth of 6.95 m) close to former CWG and SEGAS structures at the site. The extreme number of processes used throughout the years of

operation at site S7 mean it is difficult to predict the source of sample 12, though it is likely to be associated with the CWG or SEGAS processes located nearby the sampling point.

**Site S8**    Coal tar sample 13 was obtained from an FMGP site (S8) which opened in 1885. The plant hosted a series of horizontal retorts and two CWG plants for gas production. In 1935, a gas main was constructed to allow the site to store, purify and distribute coke oven gas from a nearby coke and by-product works. This resulted in decommissioning of the horizontal retorts with the CWG plants kept purely for peak load use, thus due to the greater period of operation, coal tar sample 13 was most likely formed by CWG processes.

**Site S9**    Two separate tar samples were provided from a very large FMGP at site S9, which operated horizontal retorts, coke ovens and a CWG plant. The approximate sampling locations are identified in Figure 4.3. The first sample (no. 14) was obtained from within a former gasholder, which was later used as a tar emulsion tank; thus the sample was attributed to a CWG source. On the other hand, the second sample (no. 17) was taken from a pool of tar in an area near both the CWG plant and a large, horizontal retort house, meaning it could be associated with either process.

**Site S10**    The FMGP at site S10 operated from 1841 until 1961 as a Tully gas plant which used a combination of vertical retorts and water gas processes. It is believed that early horizontal retorts were also located on site, however, tar sample 15 is thought to have been produced by vertical retorts from the Tully gas plant. Sample 15 was obtained from a sump of tar/oil and is expected to have been exposed to significant weathering.

**Site S11**    Coal tar sample 16 was acquired from a typical, small country gasworks in operation from 1854 until 1946. The site only every employed horizontal retorts for gas production. Site plans, dated 1961, indicate that a single gasholder with a maximum capacity of 21,000 cf was employed at site S11. The tar sample was taken from a borehole (at a depth of 1.9 m) in the location of a former tar tank, associated with the simple, low temperature, horizontal retorts. The site has since been remediated and redeveloped for commercial and residential purposes.
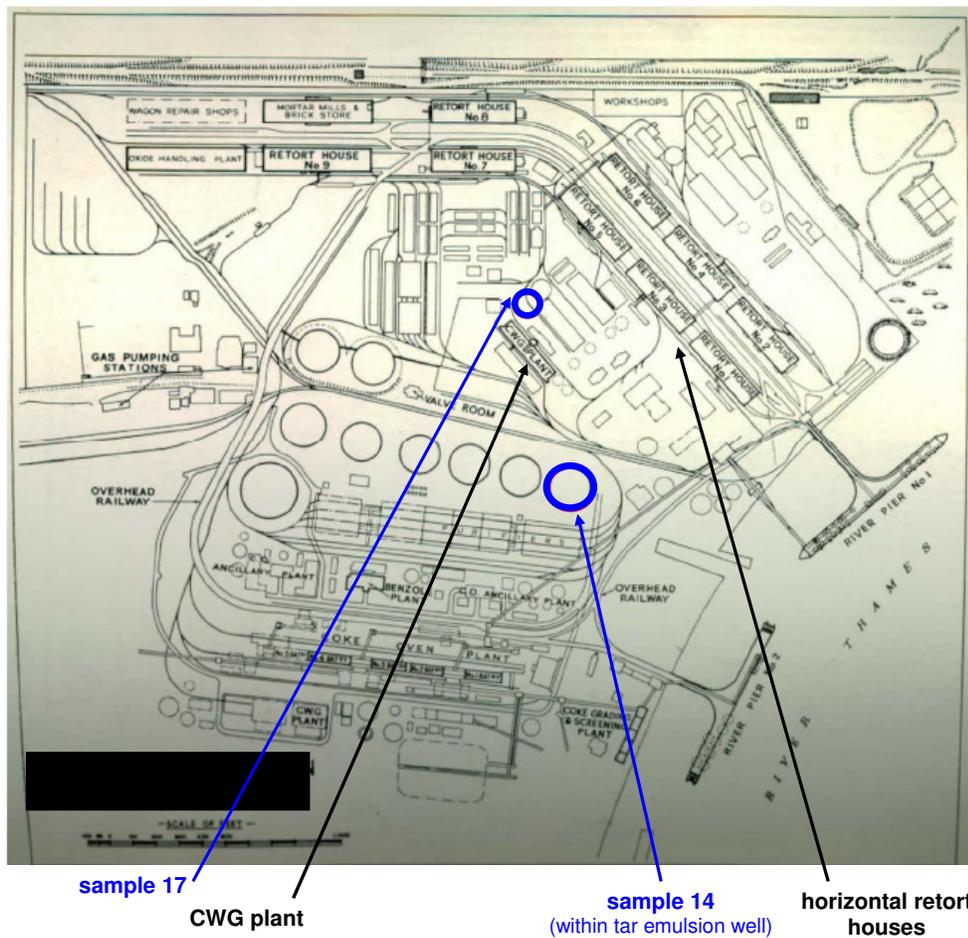
Figure 4.3: Approximate sampling locations of coal tar samples 14 and 17 at site S9. The map has been anonymised for confidentiality reasons.

**Site S12**   Site S12 is a present-day steel works in which coke ovens are still operational. Tar samples were acquired from two different types of coke oven within the site. Sample 18 was removed from a high temperature coke oven constructed in the 1970s, while sample 19 was taken from an older (circa 1930s), lower temperature coke oven. Both tar samples were freshly taken from the individual tar storage tanks on the day of production. Furthermore, an aliquot of fresh tar no. 19 was left open to the air in a fumehood for three months to provide a weathered version of the sample (labelled as W19 in Chapter 7).

**Site S13**   Initial gas production (between 1885-1920) at site S13 was likely to have been performed by advanced horizontal retorts. In 1920, a continuous vertical retort plant was constructed and expanded in 1944 to provide a CVR bench capable of gas-making at a rate of 1,613,000 cf per day. Additional vertical retorts were constructed in 1946 to boost gas production by 2,132,000 cf per day.

Historical plans of the site (circa 1960) also depict a CWG structure, therefore the acquired tar sample (no. 20) is likely to be from either vertical retorts or CWG processes, or a mixture of both. The site also operated a reforming plant using naphtha, though this is unlikely to have produced significant amounts of tar, naphtha leakages may have dissolved into any tar with which it came into contact. Tar sample 20 was provided in five separate containers, therefore, to ensure the sampling method did not affect the results, three containers were chosen at random and analysed individually (labelled as 20t1, 20t3 and 20t5). Site S13 was part of a blind study trial, the site processes detailed in this paragraph were not disclosed until analysis and data processing of the samples was completed; further details of the blind study are provided in Chapter 8 of this thesis.

**Site S14**   Samples 21 and 22 from site S14 were also originally intended to be included in the blind study, however, it was disclosed by Parsons Brinckerhoff that these samples represent a very rare form of coal tar which would be beneficial to include in the initial dataset. Site S14 was a small-scale (only one or two retort benches) gasworks plant which closed by 1870, making the tar samples unique as they could only have been formed by early, low temperature horizontal retorts. After 1870, the surrounding area was supplied with gas from a nearby gasworks, therefore the site was not used for any further coal carbonisation processes.

Samples 21 and 22 were taken from the base of a gasholder, at depths of 1.8 m and 2.3 m respectively,

and are likely to have been subjected to significant weathering. These samples represent the lowest temperature coal tars examined in this study. As this type of coal tar is rare to find (and thus very interesting for comparison with the initial coal tar library) it was decided that site S14 should not be included in the blind study as part of Chapter 8.

**Site S15**   Site S15 is thought to have been constructed between 1896 and 1915 on the site of a former mill. The site operated low temperature horizontal retorts in the beginning, producing 90.4 million cf of gas in 1915. In 1916, continuous vertical retorts were commissioned to allow production of 378,000 cf of gas per day. Further expansion of the site, including construction of eight more CVR benches occurred in 1938 to provide an additional 500,000 cf of gas per day. In 1972, the conversion process to natural gas began and production at the old gasworks ceased.

A single tar sample (labelled as 23) was attained from a borehole within a former gasholder at the site. The major process of gas production throughout the years of operation was vertical retort carbonisation, however, a CWG plant was also constructed at site S15 in the 1950s, thus sample 23 may incorporate traces of CWG tar. Site S15 was part of a blind study trial, the site processes detailed in this paragraph were not disclosed until analysis and data processing of the samples was completed.

### 4.1.2   Physical Appearance of Coal Tars

The physical appearance of each coal tar sample was recorded and the details are summarised in Table 4.2. The data in Table 4.2 shows no clear pattern between manufacturing process and physical appearance of tar samples, however vague trends do exist. Retort carbonisation tars tend to be darker in colour and more viscose, especially those formed by high temperature horizontal retorts. This trend agrees with findings by Pryde [1934], that low temperature carbonisation produced brown, oily tars while high temperature carbonisation gave black, viscose tars. In addition, the creosote sample and tars produced by CWG processes all had low viscosity and a slight red colouring. Recording a physical description of the tar could prove a useful first step in determining the manufacturing process(es) used to produce them. On the other hand, FMGP wastes will typically have been exposed to environmental processes within the subsurface for decades, if not centuries. The variety of possible processes could affect the physical characteristics of the tar samples. For example, dissolution or evaporation of smaller molecules may result in a heavier, residual DNAPL.

89

Table 4.1: Summary of manufacturing processes employed at FMGP sites (S1-S15) investigated in this study

| DNAPL label | Site label[a] | Process(es) employed | Years of Operation |
|---|---|---|---|
| 1-6 | 1 | Vertical retorts; potential traces of CWG, oil reforming and early horizontal retort tar | 1836 - 1971 |
| 7 | 2 | Horizontal retorts | 1856 - 1969 |
| 8 | 3 | Horizontal retorts | 1886 - 1971 |
| 9 | 4 | Horizontal (early, low temperature) and vertical retorts plus CWG plant | ceased production by 1953 |
| 10 | 5 | Horizontal retorts | 1849 - 1981 |
| 11 | 6 | Wood preservation site - distillation of coal tar for creosote oil | unknown |
| 12 | 7 | CWG plant | 1854 - unknown |
| 13 | 8 | CWG plant | 1885 - unknown |
| 14 and 17 | 9 | Horizontal retorts; potential traces of CWG and coke oven tar | unknown |
| 15 | 10 | Tully gas plant; combination of vertical retorts and water gas | 1841 - 1961 |
| 16 | 11 | Horizontal retorts (early, low temperature) | 1854 - 1946 |
| 18 | 12 | Coke ovens | 1970's - present day |
| 19 | 12 | Coke ovens | 1930's - present day |
| 20[b] | 13 | Vertical retorts; potential traces of CWG tar | 1885 - unknown |
| 21 and 22 | 14 | Horizontal retorts (early, low temperature) | ceased production by 1870 |
| 23[b] | 15 | Vertical retorts | 1915 - 1979 |

[a] Site labels correspond to those presented in Figure 4.1

[b] Sample used as part of a blind study, site details only disclosed by Parsons Brinckerhoff after analysis and data processing were completed.

Table 4.2: Physical appearance of coal tar samples 1-23 investigated in this study.

| Sample | Colour | Viscosity | Additional Comments | Dominant MGP Process(es)[a] |
|---|---|---|---|---|
| 1 | dark brown | high | gritty | VR |
| 2 | red/brown | high | gritty | VR |
| 3 | dark brown | high | gritty | VR |
| 4 | dark brown/black | low | fluid at top, but sludge-like at the base of the vial | VR |
| 5 | dark brown | very low | fluid at top, but sludge-like at the base of the vial | VR |
| 6 | dark brown/black | very low | fluid at top, but sludge-like at the base of the vial | VR |
| 7 | black | very high | sticky, highly viscose tar | HR |
| 8 | dark brown/black | high | high viscosity tar | HR |
| 9 | black | low | low viscosity on top, sticky tar at base | HR |
| 10 | black | low | low viscosity brown oil on top, tar at base | HR |
| 11 | dark red/brown | very low | easily poured liquid | CR |
| 12 | orange/red/light brown | very low | easily poured, clear liquid | CWG |
| 13 | dark red/brown | very low | easily poured liquid | CWG |
| 14 | red/brown | low | easily poured, smooth liquid | HR/CWG |
| 15 | dark brown | very high | very gritty | Tully (VR) |
| 16 | black | very high | tar mixed with some soil and fine gravel | HR |
| 17 | black | very high | very sticky, viscose tar | HR |
| 18 | black | very high | very sticky, viscose tar | CO |
| 19 | black | very high | very sticky, viscose tar | CO |
| 20 | black | very high | sticky, viscose tar (same for all 5 tubs of sample 20) | blind study |
| 21 | dark brown | low | gravel coated in thin layer of tar | HR (low temperature) |
| 22 | dark brown | low | gravel coated in thin layer of tar | HR (low temperature) |
| 23 | brown | very low | low viscosity tar; collected with a water sample | blind study |

[a] VR=vertical retorts, HR=horizontal retorts, CWG=carburetted water gas, CO=coke ovens and CR=creosote.

## 4.2    Standards

All solvents used (acetone, n-hexane, dichloromethane, isooctane) were of analytical grade, purchased from Fisher Scientific (Loughborough, U.K) and used without further purification. Deuterated PAHs were obtained from Isotec™, Sigma Aldrich (Gillingham, UK). All alkanes, PAHs and alkylated without further purification. Deuterated PAHs were obtained from Inaphthalene standards were also purchased from Sigma Aldrich. Octafluoronaphthalene (the internal standard for CSIA) was also obtained from Sigma Aldrich. A certified isotopic reference standard composed of a mixture of n-alkanes (C16-C30) was purchased from the Department of Geological Sciences at Indiana University (Bloomington, ID, USA).

## 4.3    Sample Preparation Techniques

### 4.3.1    Accelerated Solvent Extraction (ASE)

Extraction was performed using an ASE 350 Accelerated Solvent Extraction system (Dionex, Camberley, UK) equipped with 5, 10, 34 and 100 mL stainless steel extraction cells.

Various ASE methods were developed to meet the requirements of different experiments. For example, samples to be analysed by CSIA required fractionation while those to be analysed by GCxGC TOFMS could be analysed as a single extract. The specific instrumental parameters used for each experiment are described in the Materials and Methods sections of the relevant results chapters.

For all extractions, ground mixtures of coal tar, diatomaceous earth (an inert filtration agent) and sodium sulfate (to remove residual water) were prepared in approximately a 1:1:1 ratio. Deuterated surrogates (D8-naphthalene, D10-fluorene, D10-fluoranthene and D12-chrysene prepared as a 2000 $\mu$g/mL stock solution in dichloromethane) were added at this stage to monitor any loss of analytes from the beginning of the sample preparation process. Deuterated surrogates were not added to those extracts intended for isotope analyses as deuterated compounds are detrimental to the instrument. ASE extraction cells (5 or 10 mL) were prepared by lining the lower lid with two filter papers to collect unwanted particulates, then filling the cell with 3 g of silica gel (deactivated by

10% w/w with deionised water). A portion of the coal tar mixture (between 0.3-1.5g) was then added and the remaining cell volume packed with diatomaceous earth. The cell was then closed tightly and placed on the ASE sample carousel ready for extraction to commence.

All coal tar samples were extracted in duplicate. Extraction cells and vials were cleaned thoroughly and rinsed with acetone after each use to prevent cross contamination. Rinse cycles were also performed between samples and procedural blanks (blank extraction cells) were run at the start, middle and end of extraction sequences to ensure no crossover contamination occurred between samples.

## 4.3.2   Sample Evaporation

To obtain concentrations suitable for quantitative analysis, sample extracts were concentrated to approximately 1 mL using a Büchi Syncore®Analyst (Oldham, UK). The Büchi Syncore®Analyst is a modular system consisting of a freezer unit, V700 vacuum pump connected to a V-855 vacuum controller, a heated sample platform and a condenser unit. The freezer unit and heating racks were switched on 15 minutes prior to use to allow the required temperatures to be achieved (between -1 and -5 °C for the freezer unit).

For samples using hexane as a solvent, the temperature of the sample rack was maintained at 60 °C with the lid heated to 70 °C. The vacuum pump controller operated a three step program as follows: 500 mbar to 300 mbar in 3 minutes, 300 to 200 mbar in 2 minutes with the final pressure of 200 mbar maintained for a further 2 minutes. This gave a total evaporation time of 7 minutes. When a mixture of hexane and toluene was used as the extraction solvent the same vacuum pump settings were employed, but the sample rack and cover temperatures were increased to 70 °C and 80 °C respectively.

After evaporation, samples were re-diluted with hexane to an exact volume (generally 2-3 mL). A 1 mL aliquot was then removed to an autosampler vial and spiked with internal standard prior to analysis. Extracts were stored at -80 °C between analyses in sealed vials to prevent loss of analytes through evaporation.

Annotated photographs of the ASE 350 and Büchi Syncore® Analyst are presented in Appendix A.

## 4.4    Analytical Instrumentation

### 4.4.1    GC-MS Analyses

A Thermo Scientific (Hertfordshire, U.K.) Trace Ultra GC fitted with a DSQII quadrupole mass spectrometer and Triplus autosampler was used for all GC-MS analyses. Helium (BOC Ltd., 99.999 % purity) was used as the carrier gas at a flow rate of 1 mL/min. A J&W Scientific DB-5 column with dimensions 30 m x 0.25 mm id x 0.25 μm film thickness was used for all GC-MS analyses.

Data processing was performed using the Xcalibur$^{®}$ software (Thermo Scientific Corporation, Massachussetts, USA). The mass spectra of unknown organic compounds were compared against those within the National Institute of Standards and Technology (NIST) Mass Spectral Library (Gaithersburg, MD, USA).

### 4.4.2    GCxGC TOFMS Analyses

All GCxGC TOFMS analyses were performed using a Leco (St Joseph, Michigan) time-of-flight mass spectrometer, model Pegasus 4D, connected to an Agilent 7890A gas chromatograph equipped with a Leco cryogenic modulator. Sample injections were performed using an MPS2 twister autosampler (Gerstel, GmbH & Co., Germany). Helium (BOC Ltd., 99.999 % purity) was used as the carrier gas at a flow rate of 1 mL/min.

The GCxGC TOFMS system was optimised for the analysis of coal tars; the process is described thoroughly in Chapter 6. Each extract injected onto the GCxGC TOFMS contained four deuterated surrogates to monitor extraction efficiency and an internal standard (D10-phenanthrene) to monitor instrumental performance.

Data processing was performed using the ChromaTOF$^{®}$ software (version 4.22, Leco). The mass spectra of unknown organic compounds were compared against those within the National Institute of Standards and Technology (NIST) Mass Spectral Library (Gaithersburg, MD, USA).

**Calibration**

Alkylated naphthalenes were identified in the coal tar extracts using individually prepared 200 $\mu$g/mL in dichloromethane standards of 1- and 2-methyl naphthalene (herein referred to as the C1

alkyl naphthalenes), 1- and 2-ethyl naphthalene, 1,2-dimethyl naphthalene, 1,3-dimethyl naphthalene, 1,4-dimethyl naphthalene, 1,5-dimethyl naphthalene, 1,6-dimethyl naphthalene, 1,7-dimethyl naphthalene, 1,8-dimethyl naphthalene, 2,3-dimethyl naphthalene, 2,6-dimethyl naphthalene and 2,7-dimethyl naphthalene (herein referred to as the C2 alkyl naphthalenes).

Target analytes in the coal tar extracts were quantified by GC-MS and GCxGC TOFMS using calibration mixtures containing 16 PAHs, priority pollutants as listed by the US EPA. The 16 PAHs were purchased as a pre-prepared 2000 $\mu$g/mL stock solution in benzene:dichloromethane (1:1). A 2000 $\mu$g/mL stock surrogate solution containing deuterated PAHs (D8-naphthalene, D10-fluorene, D10-fluoranthene and D12-chrysene) was prepared and included in the calibration mixture. Seven calibration standards containing the PAHs and surrogates were prepared within the concentration range of 2.5-500 $\mu$g/mL, with each millilitre of sample spiked with 75 $\mu$L of a 2000 $\mu$g/mL stock solution of D10-phenanthrene as an internal standard. An example set of calibration curves are presented in Figure A.9 of Appendix A. Calibration curves are generally required to have a correlation ($r^2$) of at least 0.99 to be acceptable for quantification [Hunkeler et al., 2008].

### 4.4.3   GC-C-IRMS Analyses

The system used for compound specific carbon isotope analysis comprised of a Trace GC, GC Isolink and Conflo IV interfaces and a Delta V advantage isotope ratio mass spectrometer (all Thermo Fisher Scientific, MA, USA). A schematic of the system is presented in Figure 4.4.

The GC Isolink interface consists of both a reactor for high temperature thermolysis for hydrogen analysis and a combustion reactor for carbon analysis, though only carbon analyses were performed in this study. A simple valve system enables the switching of one mode to the other. As described in the previous chapter, the combustion reactor is a 0.5mm ID ceramic tube filled with nickel oxide and copper oxide wires (labelled as #5 in Figure 4.4). The combustion reactor was maintained at 1020°C and reoxidised frequently (approximately every 50 analytical runs) using a stream of oxygen gas. The flow of reference gases and the introduction of the GC flow into the IRMS is controlled by the Conflo IV interface. An open-split (labelled # 7 on Figure 4.4) was used to control variations in pressure prior to introduction of gases into the ion source. Helium (BOC Ltd., 99.999 % purity) was used as the carrier gas at a flow rate of 1 mL/min. Data processing was performed using the Isodat® software (Thermo Scientific).
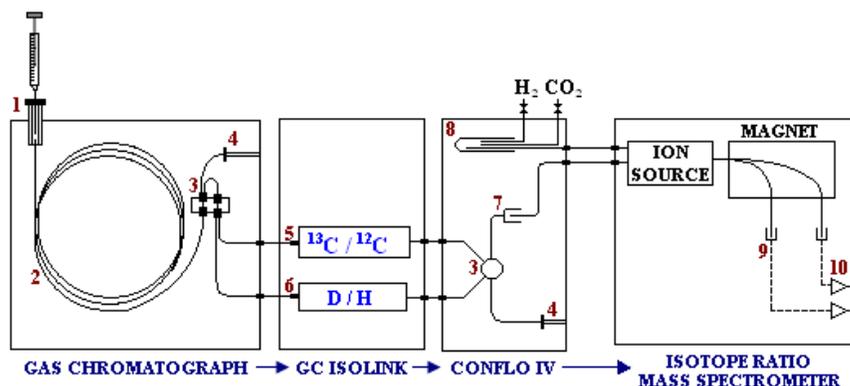
Figure 4.4: Schematic of the Thermo Scientific GC-C-IRMS system (adapted from ThermoScientific, 2008).

Octafluoronaphthalene (purchased from Sigma Aldrich, Gillingham, UK) was added as an internal standard (150 $\mu$L spike using a 2000 $\mu$g/mL stock solution) for isotope analyses as deuterated compounds are detrimental to the combustion reactors. A certified isotopic reference material containing a mix of C16-C30 n-alkanes was obtained from the Department of Geological Sciences at Indiana University (Indiana, USA) to calibrate the carbon dioxide reference gas prior to beginning sample analysis.

Annotated photographs of the GC-MS, GCxGC TOFMS and the GC-C-IRMS are presented in Appendix A.

## 4.5   Information Technology Software

**Chromatographic Data Processing**   GCxGC data processing was performed using the ChromaTOF® software (version 4.22, Leco). The software was programmed to automatically detect peaks within each chromatogram using preset peak requirements, such as peak width and minimum signal to noise ratio (S/N). The minimum signal to noise ratio was set to 10 for data processing of all GCxGC analyses performed in this thesis. The software then compares the the detected peak against the NIST mass spectral database to provide a list of probable molecular assignments.

**Data Preprocessing** Raw chromatographic data was collated and normalised against the internal standard peak using Microsoft Excel software (version 11.8). Preprocessing transformations such as root, reciprocal and logarithmic transformations were then performed.

**Statistical Packages** A combination of two different statistical packages were used throughout this study; Minitab® (version 16, Minitab Ltd., Coventry) and Matlab® (R2011a, version 7.12, Mathworks Inc., Massachusetts).

## 4.6 General Statistics

Statistical analysis of results was a crucial element of this study. This section will provide an overview of basic statistics used frequently within the research, while more specific and complex statistical techniques will be described in the relevant results chapters.

Analytical results were compared and described using the average value of repeated measurements and the spread of measurements around the average value.

**Mean** The average value of a set of numerical values is described by the mean, x, and is given by Equation 4.6.1.

$$\bar{x} = \frac{\sum x}{n} \tag{4.6.1}$$

where n refers to the total number of measurements.

**Standard Deviation** The spread of values around the mean was measured by the standard deviation, s. The standard deviation is described by:

$$s = \sqrt{\frac{(x - \bar{x})^2}{(n-1)}} \tag{4.6.2}$$

The standard deviation may also be expressed in the form of a percentage using the relative standard deviation (RSD), as given by Equation 4.6.3.

$$RSD = 100\frac{s}{\bar{x}} \tag{4.6.3}$$

**Confidence Interval**   An additional method of comparing the spread of analytical measurements is the confidence interval (CI). Unlike the standard deviation which compares the spread of values around the mean, the confidence interval provides a range of values which is likely to include an unknown population parameter. The confidence interval is calculated using the standard error of the mean (SEM), as shown in Equation 4.6.4.

$$\text{SEM} = \frac{s}{\sqrt{n}} \tag{4.6.4}$$

Confidence intervals are generally calculated using a 95 % confidence level. When the data is assumed to be normally distributed, such a confidence level covers 95% of the area within a normal curve. Thus, the probability of the true mean of the population to be within $\bar{x} \pm 1.96\frac{s}{\sqrt{n}}$ is 95%.

In terms of isotopic analysis, confidence intervals are used to monitor instrumental precision. The difference between the mean values of two datasets can be compared, for example, to monitor day-to-day variation in instrumental performance.

**Student t test**    The unpaired two-sample t-test was used in this thesis to compare two independent samples. A t-value is obtained which measures the closeness of the two means within each population and is calculated by:

$$t = \frac{|\bar{x} - \bar{y}|}{\sqrt{\left(\frac{s_x^2}{n_x} + \frac{s_y^2}{n_y}\right)}} \tag{4.6.5}$$

Where $s_x$ and $s_y$ are the standard deviations of the two samples and $n_x$ and $n_y$ are the numbers of replicates in the two samples.

The null hypothesis used in this thesis was that both arithmetic means were equal and the t-tests used were two-tailed, thus the null hypothesis was rejected if the t-values were under either tail of

the normal statistical distribution curve. A 95% confidence level was used for all t-tests performed in this thesis, meaning that the tail of the normal distribution curve was 2.5% on each side.

## 4.7    Quality Control Procedures

The US EPA describe a quality assurance system as being a process to ensure that all data "be scientifically valid, defensible, and of known precision and accuracy" [USEPA, 1992], which, as explained earlier, is imperative for environmental forensic applications. To ensure the accuracy of the analytical data produced, stringent quality control (QC) measures were employed:

- Operator training on all analytical equipment

- Procedural blanks and reagent blanks

- Calibration with standards

- Duplicate extraction and analysis of samples

- Glassware cleaned thoroughly between uses and rinsed with acetone

- Gloves changed between handling different coal tar samples

### 4.7.1    Isotopic Quality Control

The complex nature of isotope analysis requires the application of additional quality control procedures to ensure accurate and precise isotopic values are obtained.

**Calibration**     In isotopic analysis, $\delta$ values are calculated by comparing the isotopic ratio of a sample to that of an international standard. It is not reasonable to do this in practice, therefore, a reference gas calibrated against the international reference is generally used as a replacement. For the purpose of this study, isotopically labelled $CO_2$ was purchased from Air Liquide (Paris, France). The same cylinder of carbon dioxide was used as the reference gas for all isotopic measurements reported in this thesis.

**Stability**   The stability of the isotope ratio mass spectrometer was monitored using a stability method as described in the previous chapter. A stability run consisted of the isotopic analysis of ten identical reference gas pulses. The repeatability of the instrument was assessed by measuring the standard deviation of the $\delta$ values. Stability was considered achieved when the standard deviation of the $\delta$ values was less than 0.1 ‰(as recommended by the instrument manufacturer).

**Linearity**   As mentioned in the previous chapter, the linearity of the IRMS must be monitored to ensure $\delta$ values are independent of peak intensity. When a series of reference gas pulses with varying intensity are measured, the IRMS is considered linear if the $\delta$ values have a standard deviation of less than 0.2 ‰ (as recommended by the instrument manufacturer).

## 4.8   Summary

This chapter has outlined the specific analytical instruments, software, samples and standards used throughout this project; more detailed descriptions of their use can be found in individual Materials and Methods sections within the relevant results chapters. The next Chapter presents the development of a novel automated fractionation method for coal tars prior to compound specific isotope analysis of the aromatic fractions for potential source apportionment.

# Chapter 5

# Isotopic Analysis of Coal Tar

## 5.1 Introduction

Traditional approaches to source apportionment are limited to identification and comparison of chemical components in contaminated samples, with those found in potential source samples. Recently, a more advanced analytical method capable of comparing the stable isotopic composition of chemicals has been developed. Theoretically, the isotopic composition of a contaminant should relate to the isotopic composition of the original raw material used in its production. For example, fresh petroleum distillates produced using the same batch of crude oil and by the same distillation process should theoretically have the same isotopic signature. For this reason, isotopic analysis has recently become a major tool for source apportionment of environmental contamination.

### 5.1.1 Stable Isotopes

Within the nucleus of an atom are positive particles (protons) and neutral particles (neutrons), the sum of which represents the total mass of the atom. Each element in the periodic table can be expressed using the notation, $^{A}_{Z}X$, where A and Z represent the atomic mass and atomic number of element X respectively. The atomic mass is the sum of protons and neutrons, while the atomic number defines the element by representing the number of protons. An isotope is an atom containing

the same number of protons but a different number of neutrons, and thus a different atomic mass [Hoefs, 1980]. Isotopes can be defined as either stable or unstable (radioactive) based on whether or not they decay over time [Hoefs, 1980]. Radioactive isotopes will decay with time into more stable forms, while stable isotopes are already in a stable state, thus do not decay.

The most commonly investigated stable isotopes are elements frequently found in organic contaminants; $^{12}$C and $^{13}$C, $^{1}$H and $^{2}$H, $^{14}$N and $^{15}$N, $^{16}$O and $^{18}$O and $^{35}$Cl and $^{37}$Cl. The lighter stable isotope (i.e. smaller atomic mass) is always the most abundant in nature [Carter et al., 2005]. Generally, a single isotope is dominant in nature, with others only found in trace amounts [Hoefs, 1980]. For example, the relative abundances of the two naturally occurring stable isotopes of carbon, $^{12}$C and $^{13}$C, are 98.89% and 1.11% respectively.

### 5.1.2   Isotope ratios

The isotopic composition of an element is expressed as a ratio between the abundances of two stable isotopes. The ratios are expressed in delta ($\delta$) notation, in units per mille (‰), relative to the isotopic composition of an internationally agreed standard (Equation 5.1.1). Stable isotopes are expressed in this form as the differences between samples and standard are so small at natural abundance levels that they may only differ within the third or fourth decimal place.

$$\delta^{13}\mathrm{C}\,(\text{‰}) = \left( \frac{\mathrm{R_{sample}} - \mathrm{R_{standard}}}{\mathrm{R_{standard}}} \right) \mathrm{x}1000 \qquad (5.1.1)$$

where R represents the ratio of heavy to light isotopes.

Initially, the primary standard for carbon isotope analyses is a calcium carbonate known as Pee Dee Belemnite (PDB) with a $\delta$ value set to 0.0 ‰ to provide a baseline for isotopic measurement. However, supplies of PDB have now been exhausted and thus replaced by other reference standards calibrated against PDB, such as NBS18 [Hoefs, 1980]. When isotopic ratios are reported relative to such reference standards they are said to be "versus VPDB", in which V stands for Vienna, where the supplier of the standards, the International Atomic Energy Agency (IAEA), is based [Carter et al., 2005].

### 5.1.3    Isotopic Fractionation Processes

Isotope ratios are not constant; they may be altered by a number of chemical and biological processes. This change in isotopic composition is known as isotopic fractionation and occurs through two main mechanisms; the kinetic isotope effect and the equilibrium isotope effect [Meier Augenstein, 1999].

**Kinetic Isotope Effect**    During chemical reactions, the isotopic composition of a molecule will be affected by bond formation and breakage processes. Lighter isotopes form weaker bonds than heavier isotopes and consequently are more reactive. As a result, any molecules containing the heavier isotope will be slower to react, causing the reactant to become isotopically heavy while the product will become enriched in the lighter isotope. The kinetic isotope effect (KIE) describes the ratio of rate constants for reaction of light and heavy isotopes [Hoefs, 1980]. The variation in isotope signature can be used to monitor natural attenuation of compounds when the kinetic isotope effect is evident within metabolic degradation pathways [Hall et al., 1999]. However, for large molecules, such as PAHs, the effect is often diluted due to the site of reaction on the molecule being only a single carbon atom. Thus, during compound specific isotope analysis, where the carbon isotope ratio represents an average value for all carbon atoms in the entire molecule, the overall effect can often be very small for large molecules.

**Equilibrium Isotope Effect**    Unlike the kinetic isotope effect, the equilibrium isotope effect does not involve changes in molecular structure via a chemical reaction. The process involves the replacement of one isotope with another in a form of "isotope exchange". In the context of environmental forensics, non-degradative processes such as evaporation and dissolution are likely to be affected by isotope exchange potentially leading to differences in the isotopic signature of a contaminant when it is examined in different forms (e.g. air monitoring versus groundwater samples) [Gray et al., 2002].

### 5.1.4    Stable Isotope Analysis

In the environment, stable isotope analysis can be used to link a contaminant to a specific source. There are currently three main forms of stable isotope analysis; bulk, compound-specific and

position-specific.

**Bulk Stable Isotope Analysis**   Bulk stable isotope analysis (BSIA) measures the isotopic composition of an element within an entire sample. The isotope ratio of each molecule (containing the investigated element) present in the sample is averaged. The technique uses elemental analysis in conjunction with isotope ratio mass spectrometry (EA-IRMS) to convert the sample into simple gases prior to isotopic measurements [Benson et al., 2006].

**Compound Specific Isotope Analysis**   Compound specific isotope analysis (CSIA) has a distinct advantage over BSIA in that it is capable of measuring the isotope ratios of individual molecules within a mixture, resulting in the decreased use of BSIA in recent years [Philp, 2007, Schmidt et al., 2004, Sessions, 2006]. Generally, the technique uses gas chromatographic separation prior to quantitative conversion of the eluents into simple gases and isotope measurement, in a process known as gas chromatography-combustion interfaced-isotope ratio mass spectrometry (GC-C-IRMS). A further development in isotope technology known as position specific isotope analysis (PSIA) involves an enhanced level of specificity over CSIA, where the isotope ratios of particular atoms (or functional groups) within a molecule are targeted [Gauchotte et al., 2009]. The form of isotope analysis used within this study was CSIA, hence its application within environmental forensics will be discussed in more detail in the following sections. A full theoretical description of the technique can be found in Chapter 3.

### 5.1.5   Source Apportionment by CSIA

Source identification of contaminants, such as coal tars, by traditional tiered approaches to chemical fingerprinting can be challenging. The composition of coal tars from different FMGP sites may appear very similar in GC-FID and GC-MS chromatograms, thus provide inconclusive results. Stable isotope analysis was applied to environmental forensics after initial success in the field of petroleum geochemistry [Slater, 2003, Carter et al., 2005, Philp, 2007]. Atmospheric carbon dioxide plays an important role in the variation of carbon isotope signatures of fossil fuels. During photosynthesis, atmospheric carbon dioxide enters the metabolic cycle of plants and becomes incorporated into the plant material. Fossil fuels, such as coal and crude oil, are produced over many years when
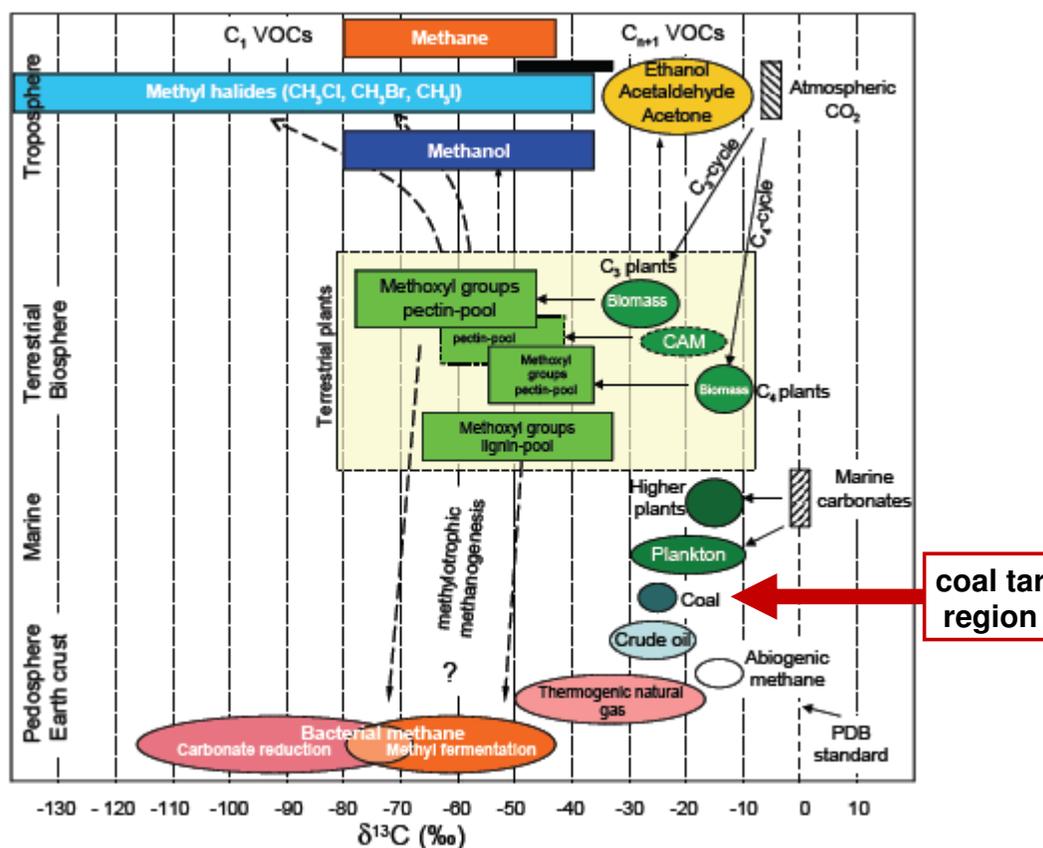
Figure 5.1: Natural variations in carbon isotope values within numerous sources. The expected region of coal tar variation is highlighted by the arrow (adapted from Keppler et al., 2004).

organic matter is compacted in layers in the Earth's crust and exposed to high temperature and pressure [Hoefs, 1980]. Therefore, any regional variations in atmospheric carbon dioxide are incorporated into the fuel. Coals will retain an isotopic signature based on the geological conditions of formation. Refined petroleum products also demonstrate this phenomenon to an extent, but will be further influenced by isotopic fractionation during distillation processes. Natural variations in carbon isotope signature are presented in Figure 5.1.

### 5.1.6   CSIA in Environmental Forensics

As previously mentioned, GC-FID and GC-MS can often prove inconclusive when investigating similar environmental samples. The incorporation of CSIA as an ancillary technique has been

investigated to aid the source identification process. The technique has been used on a variety of different samples; including gasolines [Smallwood et al., 2002, O'Sullivan and Kalin, 2008], crude oils [Mazeas and Budzinski, 2002] and refined petroleum products [Philp et al., 2002].

**Polycyclic Aromatic Hydrocarbons**

Polycyclic aromatic hydrocarbons (PAHs) can be formed by a range of anthropogenic processes. The distribution of PAHs in a sample allows classifications to be made according to the process by which they were formed; namely petrogenic or pyrogenic origin. Multiple individual sources which contribute to these main classes will have similar PAH compositions, hence it is difficult to distinguish between them. PAHs contain at least ten carbon atoms per molecule, making them well-suited for source apportionment by CSIA, as the effect of any isotopic fractionation by degradation processes will be diluted. A considerable amount of literature has been published on CSIA of numerous forms of PAH contamination.

A detailed study by O'Malley et al. [1994] presented the first use of CSIA for source apportionment in environmental forensic applications. The study evaluated the isotopic integrity of PAHs after vaporization, photolytic decomposition and microbial degradation. The results of microbial degradation studies on naphthalene showed that even after a 95% reduction in concentration, there was no significant alteration of the $\delta^{13}C$ value. In addition, O'Malley et al. [1994] showed distinct differences in $^{13}C/^{12}C$ ratios between PAH sources (including fire soots, car soots and crankcase oil) illustrating the potential for source apportionment by CSIA.

Mazeas and Budzinski [2002] also demonstrated the power of CSIA for source apportionment through analysis of oil samples collected after the Erika oil spill which occurred off the French Atlantic Coast in December 1999. The authors compared the carbon isotopic value of phenanthrenes from numerous oil residues and oil found on bird feathers obtained along the French coastline. The results, as presented in Figure 5.2, clearly support initial GC-MS evidence indicating contributions from two independent oil sources.
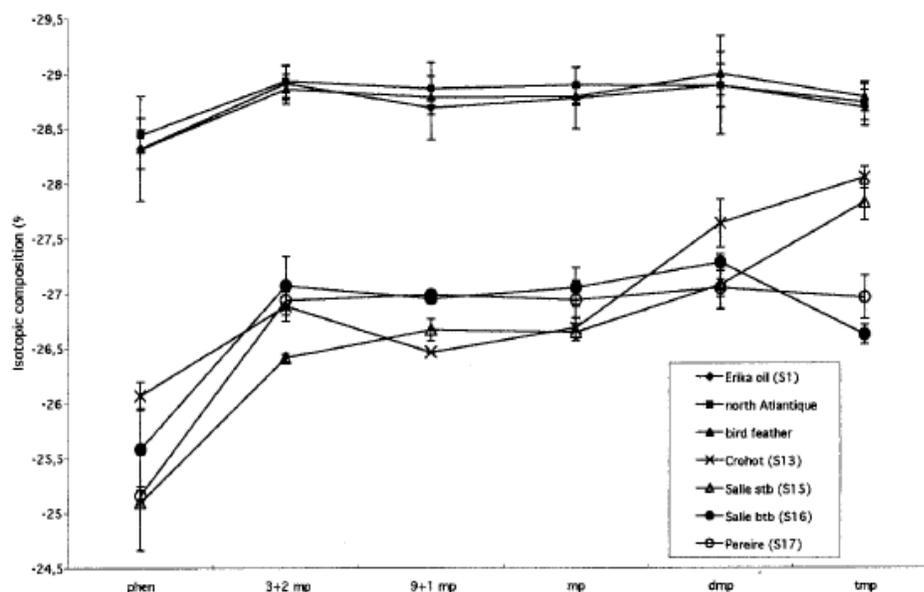
Figure 5.2: Carbon isotopic composition of phenanthrenes in oils samples collected along the Atlantic Coast of France compared with Erika oil reference samples; values represent the mean of triplicate analyses (copied from Mazeas and Budzinski, 2002).

**Coal tars**

As previously stated, environmental forensics developed due to the requirement to identify spilled oils in response to major incidents such as the Erika and Exxon Valdez spills. As such, many publications in this field focus on the analysis of crude oils and refined petroleum products. Coal tar is a ubiquitous contaminant at FMGP sites making it a global problem. The tars are likely to contain many of the same components as crude oils, nevertheless, there are few reports on isotopic composition of coal tars in the literature.

A recent report by Emsbo Mattingly and Boehm [2003] for the Electric Power Research Institute (EPRI) presented the isotopic analysis of two different tar types; coal tar from a coking works and petroleum tar from a caburetted water gas (CWG) and oil gas site. The 16 EPA PAHs in the coal tar were found to have a carbon isotope range of -25 to -27 ‰, while those in the petroleum tar were isotopically lighter with a range of -28 to -30 ‰, similar to the anticipated values for a petrogenic source. The report notes that the isotopic analysis of coal tars is a challenging process due to a number of factors. Coal tar PAHs fall within a narrow isotopic range of $\delta^{13}C$ values and many suffer performance problems, such as poor precision and sensitivity [Emsbo Mattingly and Boehm,

2003]. The report also demonstrates that it can be difficult to obtain carbon isotope signatures for severely weathered samples due to low instrumental sensitivity. These identified issues may explain the low volume of published data on CSIA of coal tars.

Nevertheless, McRae et al. [1999] have demonstrated that the carbon isotopic signatures of coal tar PAHs (especially those formed by low temperature methods) are similar to those of the parent coals; approximately -24 to -25 ‰ for UK coals. The study investigated a number of coal-derived substances produced by a range of processes. The authors highlighted a trend in carbon isotope value between the processes; $\delta^{13}C$ values became isotopically lighter going from high temperature carbonisation to gasification and combustion in sequence. McRae et al. [1999] concluded that the higher degree of ring growth by gasification processes in comparison to carbonisation results in $^{12}C$-enrichment of the product coal tar PAHs as the formation of additional bonds will preferentially incorporate $^{12}C$ rather than $^{13}C$. In addition, the authors demonstrated the potential of CSIA for source apportionment of coal-derived PAHs through analysis of soil and vegetation samples from the vicinity of an FMGP. The soil samples were obtained from the side of a main road nearby the FMGP. PAHs present in the samples exhibited $\delta^{13}C$ values in the range of -24 to -26 ‰, similar to those of the low temperature coal tars produced at the FMGP. The soil samples obtained from the side of a motorway near the FMGP exhibited isotopically light $\delta^{13}C$ values indicative of petrogenic origin, thus were likely affected by transport pollution from the motorway.

An earlier study by McRae et al. [1996] found that the $\delta^{13}C$ values of transport fuels were typically between -28 and -30 ‰, while coal-derived PAHs were in the range of -23.5 to -25 ‰. The research also demonstrated that the isotopic composition of low temperature processes produced isotopic signatures similar to those of the parent coal, while higher temperature processes gave isotopically light values. However, the study did not list the exact manufacturing processes used to produce the coal tars and certain samples were actually prepared in lab-scale pyrolysis experiments. Furthermore, the study did not examine the numerous types of gas-making processes as represented by the coal tar library collected for this thesis.

Hammer et al. [1998] investigated the isotopic composition of a coal tar distillate. The authors demonstrated that the $\delta^{13}C$ values of creosote oil PAHs lie within the narrow range of -22.7 to -25.3 ‰. However, this study focussed solely on creosote samples and did not compare the $\delta^{13}C$ values to those from other PAH sources, such as the initial coal tars distilled to produce the creosote.
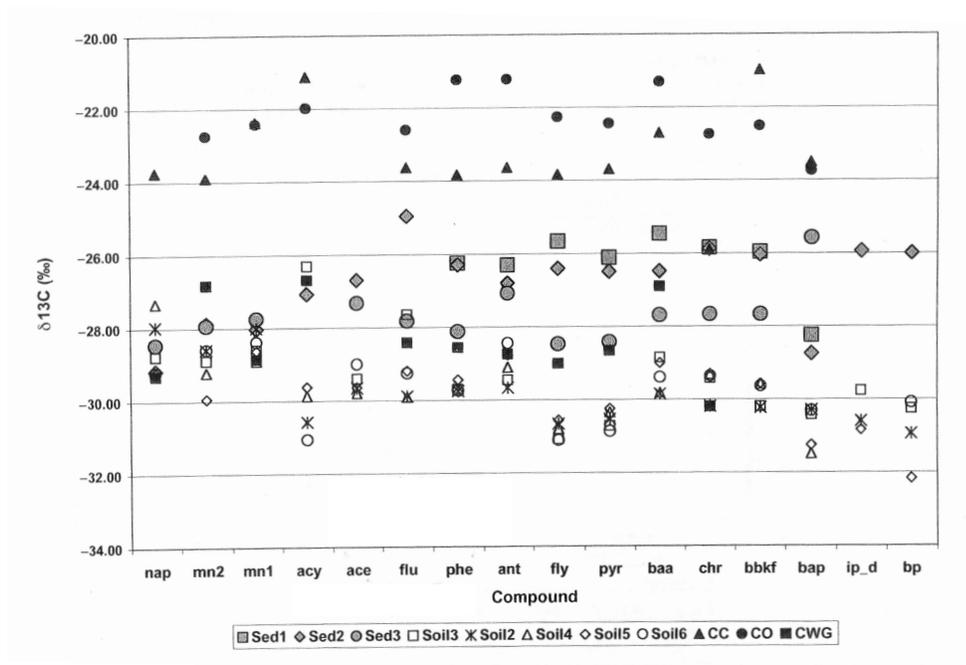
Figure 5.3: Compound specific isotope ratios of PAHs in soil and sediment samples (from a single FMGP site) compared to example tars from various other sites, where CO=coke oven, CWG = carburetted water gas and CC = coal carbonisation (copied from Saber et al., 2006).

Previous isotope studies at FMGP sites have tended to focus on the analysis of soil or sediment samples as opposed to free phase coal tars. A recent study by Saber et al. [2006] employed traditional tiered analysis of sediments near an FMGP with GC-C-IRMS as an ancillary technique. The study analysed a number of reference tars (as no tars from the site under investigation were available) from CWG, coke oven and coal carbonisation processes (Figure 5.3). CWG tars were found to be isotopically enriched in ^{13}C; which the authors attributed to the use of significant amounts of petroleum distillates at such sites [Saber et al., 2006]. The isotopic data in this investigation failed to conclusively identify the source of PAHs in the unknown soil and sediment samples. However, the samples were most similar to the CWG reference tar, while the carbonisation and coke oven tars were isotopically heavy in comparison.

So far CSIA has only been applied to coal tars produced by a limited number of manufacturing processes. A comprehensive study of coal tars from a broad range of manufacturing methods (such as in the coal tar library collated for this thesis) is necessary to discover the impact of CSIA for source apportionment at FMGP sites.

### 5.1.7   Aim and Objectives

The overall aim of this chapter was to investigate the capability of compound specific isotope analysis for source allocation of coal tar samples. The objectives set to reach this aim were:

- Optimisation of an automated sample fractionation method using accelerated solvent extraction.

- Carbon isotope analysis of the aromatic fraction of coal tar.

- Evaluation of isotopic data for source allocation of coal tars.

## 5.2   Materials and Methods

### 5.2.1   Sample Preparation

Anhydrous sodium sulfate, silica gel 60 (both purchased from Sigma Aldrich) and diatomaceous earth (Dionex, Camberley, UK) were activated for 4 hours at 450 °C prior to use. A 2000 $\mu$g/mL stock solution of octafluoronaphthalene (purchased from Sigma Aldrich) was prepared in hexane for use as an internal standard.

A dry, homogeneous mix of coal tar was prepared by grinding the coal tar (approximately 0.5 g) with sodium sulfate ($Na_2SO_4$) and diatomaceous earth (DE) in a 1:1:1 ratio. This removes any water present in the coal tar sample and results in a fine powder (rather than a tar) which can be transferred quantitatively to the extraction cells. Surrogates were not added as quantification of analytes was not deemed necessary at this stage.

### 5.2.2   Fractionation by Column Chromatography

Sample fractionation was performed by traditional column chromatography to compare the $\delta^{13}C$ values of the aromatic fraction to those obtained by analysis of extracts produced by a fully automated fractionation method.

An ASE 350 accelerated solvent extraction system (Dionex, Camberley, UK) was used to provide a sample extract in hexane ready for column fractionation, using 5 mL stainless steel extraction cells filled with approximately 1.0-1.5 g of the described coal tar mixture. The sample was extracted at 150 °C and 10 MPa pressure. The resulting extract was concentrated to approximately 1 mL using a Büchi Syncore® Analyst (Oldham, UK).

For the column fractionation procedure, a 50 mL burette was plugged with cotton wool and filled with 20 g of silica gel 60. The top of the column was also plugged with cotton wool to ensure the silica gel was firmly compacted within the burette. The column was conditioned using 30 mL of hexane prior to addition of the sample extract. The first fraction was eluted with hexane (30 mL), the second by 30 mL of hexane/toluene (50:50 mixture) and the final fraction eluted with toluene (40 mL). The fractions were all concentrated to 1 mL using a Büchi Syncore® Analyst and spiked with 150 $\mu$L of octafluoronaphthalene (from a 2000 $\mu$g/mL stock solution in hexane) as an internal standard prior to analysis.

### 5.2.3   Development of automated coal tar fractionation

An automated fractionation method was developed for the ASE 350 accelerated solvent extraction system (Dionex, Camberley, UK) equipped with 10 mL stainless steel extraction cells. Larger cells (10 mL rather than 5 mL) were used to accommodate the inclusion of silica gel for in-cell sample cleanup. The extraction cells were lined with 2 filter papers (to ensure unwanted particulate matter did not collect in the extract) and packed with 3 g silica gel 60 (10% deactivated w/w using deionised water) to provide simultaneous sample extraction and clean up. Approximately 1.0-1.5 g of the ground DNAPL mixture was added to the extraction cell and the remaining cell volume was packed with D.E.

To allow sample fractionation, three separate ASE methods were employed to sequentially extract the same cell using solvents of increasing polarity. To obtain the first fraction, hexane (50 % cell volume) was used to extract the cell. The oven and static times were switched off to allow the solvent to flow straight through the cell and encourage only the aliphatic portion to elute. The second fraction was eluted with hexane:toluene in a 9:1 ratio (70 % cell volume). The oven temperature was maintained at 50 °C with the cells heated for 5 minutes prior to extraction. The final fraction was extracted using toluene (70% cell volume) at 100 °C (with 5 minute heating time).

The extracts were concentrated to 1 mL using a Büchi Syncore® Analyst and transferred to an autosampler vial. All extracts were spiked with 150 $\mu$L of internal standard (octafluoronaphthalene) prior to analysis.

### 5.2.4    Isotope analyses

The gas chromatograph was fitted with a 30 m DB-5 capillary column (0.25mm ID, 0.25μm film thickness) supplied by J&W Scientific. The helium flow was kept constant at 1 mL/min. The initial oven temperature was set to 55 °C and held for 2 minutes before the temperature was ramped at 5 °C/min to 320 °C, with a final temperature hold time of 12 minutes. One microlitre of sample was injected using a Triplus (Thermo Scientific) autosampler. The split ratio was varied between 20 and 80 in order to obtain a signal size higher than 0.2 V for the $m/z$ 44 ion of each of the PAHs investigated. In certain cases, naphthalene was found in such high abundance relative to the other PAHs that it required to be analysed separately using a higher split ratio. To achieve this, two different analytical runs were programmed. The first injection used a high split ratio (between 100 and 300) and isolated naphthalene by programming the backflush to close only during the retention time window in which naphthalene elutes (between 600-1100 seconds). Therefore, the remainder of the sample was sent to waste when the backflush was switched on after elution of naphthalene. The second injection was used to analyse the remaining low concentration PAHs at a lower split ratio with the backflush valve set to switch on during the retention time window of naphthalene to ensure the high concentrations did not cause damage to the reactor.

All sample isotope values were calculated relative to a standard gas (carbon dioxide) injected at the beginning of each sample run. Each analytical run was programmed to include 5 reference gas pulses prior to the sample peaks and an additional 3 reference gas pulses immediately after the last sample peak.

The analytes of interest (the 16 US EPA PAHs plus the C1 methyl naphthalenes) were identified prior to isotope analysis using a Thermo Scientific (Hertfordshire, U.K.) Trace Ultra GC fitted with a DSQII mass spectrometer and Triplus autosampler. The same column was used for both GC-MS and GC-C-IRMS; a J&W Scientific DB-5 fused silica capillary column (30 m x 0.25 mm id x 0.25 $\mu$m film thickness). All injections were of one microlitre and were carried out using a split ratio of 1:50 and injection port temperature of 230 °C. Helium was used as the carrier gas, with

a flow rate of 1 mL/min and extracts were analysed using the same oven temperature program as for GC-C-IRMS analyses. The ion source and transfer line were maintained at 200 ˚C and 320 ˚C respectively. The electron ionization voltage was set at 70 eV for all analyses.

## 5.2.5   GCxGC TOFMS analyses

All GCxGC TOFMS analyses were performed using a Leco (St. Joseph, Michigan) time of flight mass spectrometer, model Pegasus 4D, connected to an Agilent 7890A gas chromatograph equipped with a Leco thermal modulator. The TOF ion source was fixed at 200 ºC and masses between 45 and 500u were scanned at a rate of 200 spectra per second. The detector voltage was set at 1700 V and the applied electron ionization voltage was set at 70 eV. All results were collected and viewed using ChromaTOF® software (Leco, version 4.22).

All standards and extracts were analysed with the primary oven temperature programmed at 10 ˚C /min from 55 ˚C (2 min isotherm) to 110 ˚C, 3 ˚C/min to 210 ˚C, then at 8 ˚C/min to 310 ˚C (15 min isotherm). The secondary oven was programmed at 10 ˚C /min from 75 ˚C (2 min isotherm) to 130 ˚C, 3 ˚C/min to 230 ˚C, then at 8 ˚C/min to 330 ˚C (15 min isotherm). This corresponds to a 20 ˚C offset between the primary and secondary ovens. The modulator temperature was also maintained at a 20 ˚C offset relative to the primary oven and the modulation period was 6 seconds with a 1.3 second hot pulse time. The transfer line was maintained at a temperature of 340 ˚C. The injection port temperature was set to 250 ˚C using a split ratio of 1:50. One microlitre of sample was injected for each run using an MPS2 twister autosampler (Gerstel). Helium was used as the carrier gas, with a flow rate of 1.0 mL/min.

Primary and secondary columns were connected via a Thames Restek Press-tight® glass connector. Numerous column sets were evaluated during the optimisation of coal tar separation; specific column parameters are listed in Table 6.1 and will be discussed in the following sections.

## 5.2.6   Development of quality control procedures

**Isotopic Reference Material**    As previously mentioned, isotopic values are calculated by comparison with an international standard, or in practice, a reference gas of known isotopic value.

Therefore, the first step in this set of experiments was to verify the exact carbon isotope value of the carbon dioxide reference gas. Replicate analysis of a certified isotopic reference material (a mixture of C16-C30 n-alkanes obtained from Indiana University) was used to calibrate the reference gas. The results were used to correct the carbon dioxide reference gas to a value of -35.4 ‰. The same reference gas cylinder was used for the duration of the project.

The isotopic values of six replicate injections of the reference material (after correction of the reference gas) are shown in Table B.1 of Appendix B. The standard deviation was found to be less than 0.4 ‰ for all components, indicating adequate instrumental stability.

**Linearity and Stability**   As explained in Chapter 3, linearity and stability tests must be performed to monitor instrumental performance. It was decided that these tests should be performed at the start of each analytical sequence and following every 16 samples (equivalent to approximately 24 hours of analyses). The standard deviation of these tests were monitored to ensure the instrumental precision remained within the manufacturer's specifications of less than 0.02 ‰/nA for linearity and less than 0.3 ‰ for stability.

**Analysis of in-house standards**   A stock solution consisting of the 16 US EPA priority PAHs, was analysed immediately after completion of a sequence of stability/linearity tests and prior to sample analysis. Triplicate injections of the solution were performed and the isotopic values monitored to ensure results were consistent. A standard deviation of less than 0.5 ‰ was desirable, in accordance with the US EPA guidelines, throughout all isotopic measurements. The mean isotopic values of PAH in-house standards, which were injected in triplicate prior to each sample sequence over a three month period, are provided in Table B.2 of Appendix B. The standard deviation of carbon isotope values across the three month period were less than 0.56 ‰ for all PAHs analysed.

In addition, coal tar extracts were separated by the analysis of blank (hexane) samples to prevent cross contamination and each sample included an internal standard (octafluoronaphthalene) to monitor instrumental performance.

## 5.3    Results and Discussion

### 5.3.1    Optimisation of fractionation

The development of a fractionation method for isotopic analysis of coal tar was required to ensure close to baseline resolution was achieved. Baseline resolution is required to ensure that $\delta^{13}$C values are not affected by coeluting components. In tiered approaches to chemical fingerprinting the method of fractionation generally employed is silica column chromatography. However, this method is extremely time-consuming and labour-intensive. This section discusses the optimisation of a fast, automated procedure for coal tar fractionation using an accelerated solvent extraction (ASE) system. ASE has been previously shown to provide consistent and reproducible carbon isotope data when used as an extraction method, prior to silica column fractionation, for PAH contaminated soils [Graham et al., 2006]. However, the use of ASE for simultaneous sample extraction and fractionation of coal tars has not been reported in the literature.

The ASE cells were loaded with silica gel 60 to mimic the process within traditional silica columns. Three fractions were obtained by sequential extraction of a single ASE cell using solvents of increasing polarity. The difficulty with coal tar samples is the ease of dissolution of the components from the sample matrix. It was discovered that optimal separation of the aliphatic (fraction 1) and aromatic (fraction 2) classes could be obtained by restricting the duration of time in which the solvent resides in the cell. This was achieved by setting the static time in the cell to zero, meaning that the solvent flowed directly through the cell similar to the process in silica column chromatography, instead of allowing the solvent to be held within the pressurised cell for prolonged periods which would encourage the elution of PAHs within the first fraction. The oven was also switched off for the initial extraction to limit the elution of PAHs within fraction 1. Fraction 3 was shown to contain no coal tar components during development of the fractionation method. This is demonstrated by an overlay of GC-MS chromatograms for the three fractions of a coal tar, as shown in Figure 5.4. Therefore, fraction 3 was not deemed necessary in the extraction process for isotopic measurements as all coal tar constituents elute within the first two fractions.
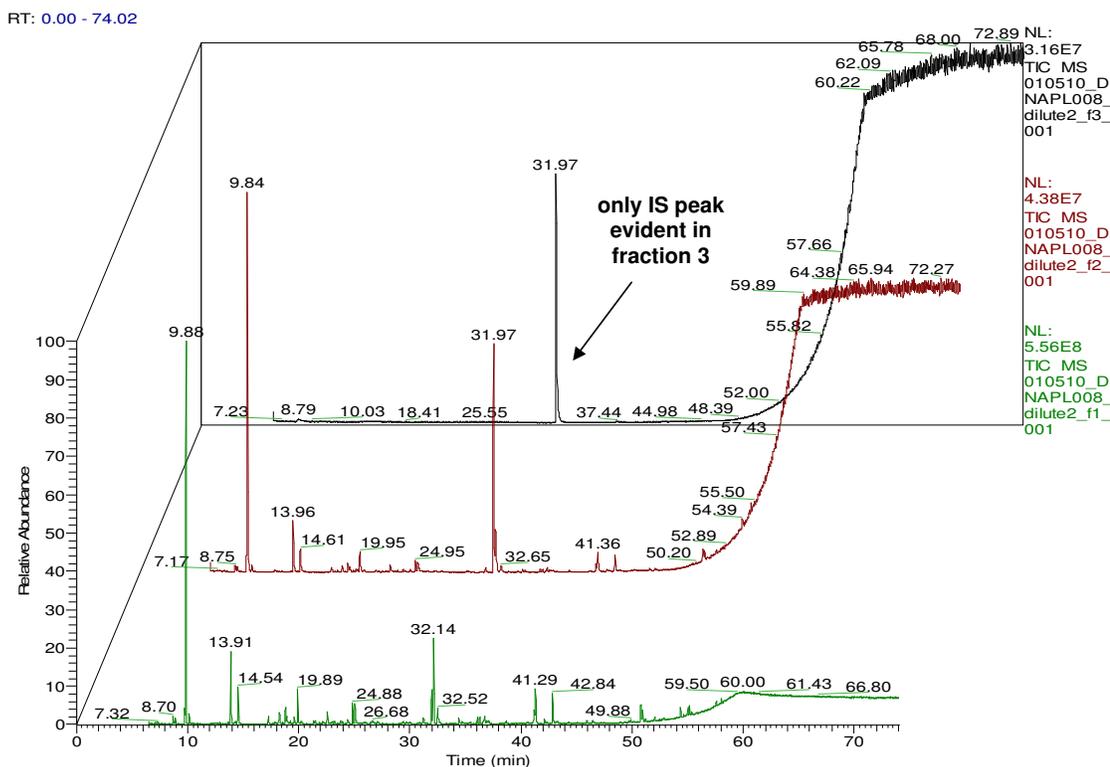
Figure 5.4: GC-MS chromatogram overlay for fractions 1-3 of coal tar sample 8. The internal standard (IS) used during the initial development of the fractionation method was a 150 $\mu$g/mL spike of D10-phenanthrene.

To test the repeatability of the ASE fractionation method, a series of three cells spiked with in-house PAH standards (16 US EPA PAHs) were extracted, using octafluoronaphthalene as a surrogate to monitor isotopic fractionation, and the extracts analysed by CSIA. The results are summarised in Table 5.1 and Figure 5.5. The results show that octafluoronaphthalene and naphthalene fully elute with the first fraction (F1), therefore isotopic values could not be obtained for these compounds for fraction 2 (F2). The octafluoronaphthalene produced poor repeatability in the extracted samples due to low peak intensity. Nevertheless, the results indicate that there is no significant difference between the PAH stock solution and the PAHs subjected to ASE fractionation. Figure 5.5 clearly shows that the PAH with the most isotopic variation ($> 1‰$) between the stock solution and the F1 and F2 extracts is anthracene. A previous study by O'Malley et al. [1994] noted an enrichment of $^{13}$C of up to 2 ‰ in anthracene due to photolytic decomposition of the compound, therefore, this may be a possible explanation for the $^{13}$C enrichment in the fractionated samples.

Table 5.1: Comparison of $\delta^{13}$C values before and after ASE fractionation of PAH-spiked extraction cells.

| PAH | Mean $\delta^{13}$C PAH value (‰) | | | | |
| | Fraction 1 (F1)[a] | Fraction 2 (F2)[a] | PAH stock[b] | $\Delta$F1[c] | $\Delta$F2[c] |
| --- | --- | --- | --- | --- | --- |
| F8-N | -20.6 | – | -21.9 | 1.3 | – |
| N | -25.4 | – | -26.2 | 0.8 | – |
| ACY | -24.9 | -24.7 | -24.7 | -0.2 | 0.0 |
| ACE | -24.4 | -24.2 | -24.5 | 0.1 | 0.3 |
| FLU | -27.0 | -27.4 | -27.7 | 0.7 | 0.3 |
| PHE | -25.3 | -25.5 | -24.9 | -0.4 | -0.6 |
| ANT | -24.7 | -24.9 | -25.9 | 1.3 | 1.1 |
| FLT | -24.8 | -25.1 | -25.3 | 0.5 | 0.1 |
| PYR | -24.6 | -25.2 | -25.5 | 1.0 | 0.4 |
| BaA | -27.1 | -26.8 | -26.4 | -0.7 | -0.4 |
| CHR | -27.1 | -26.7 | -26.9 | -0.2 | 0.2 |
| BbF | -26.8 | -26.5 | -26.3 | -0.6 | -0.2 |
| BkF | -27.0 | -26.9 | -27.3 | 0.3 | 0.4 |
| BaP | -31.3 | -31.2 | -31.5 | 0.2 | 0.3 |
| IP + DBA | -25.5 | -25.7 | -25.7 | 0.2 | 0.0 |
| BP | -30.9 | -30.1 | -30.3 | -0.6 | 0.3 |
| | | | mean | 0.7 | 0.3 |
| | | | standard deviation | 0.4 | 0.3 |

[a]Mean value of triplicate injections of 3 identical extractions (total of 9 values)
[b]Mean value of 6 replicate injections of the PAH stock solution
[c]Difference relative to initial isotope values using PAH stock solution
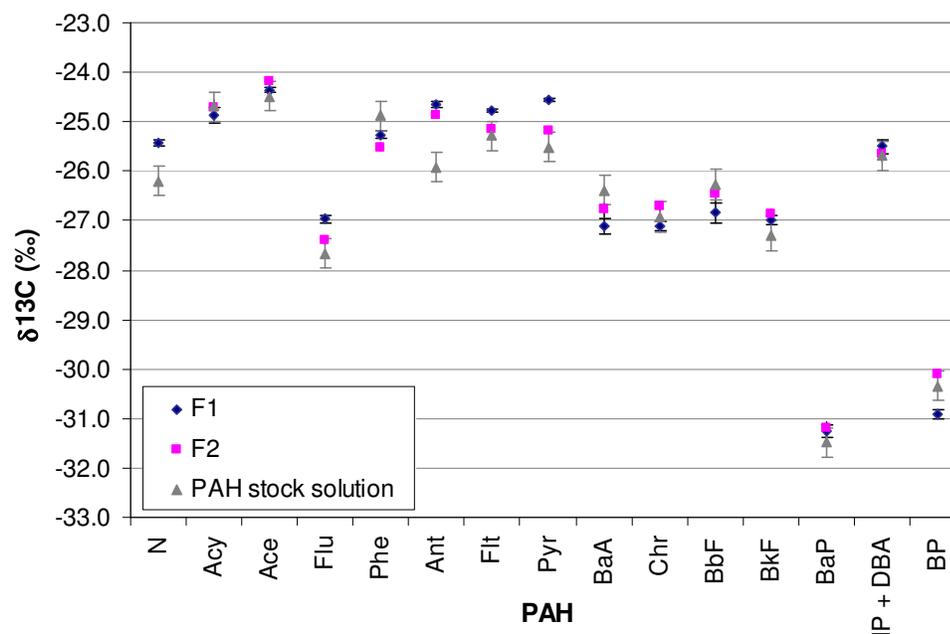
Figure 5.5: Comparison of $\delta^{13}$C values before and after ASE fractionation of a PAH standard solution. Y-error bars represent 0.3 ‰ for the PAH stock solution and a 95 % CI based on triplicate injections of 3 repeat cells for F1 and F2. Values as presented in Table 5.1.

The next step was to test the repeatability of fractionation using an actual coal tar sample. Since isotopic fractionation was not shown to be a major issue when using the ASE method (Table 5.1), the octafluoronaphthalene was used as an internal standard from this point on rather than a surrogate. Fractions of coal tar sample 7 were obtained by extraction of six replicate cells using the automated ASE fractionation method. The GC-C-IRMS procedure had to be modified slightly to account for exceptionally high levels of naphthalene in this sample. The isotopic values of naphthalene were analysed separately using a high split ratio to prevent overloading and damage to the reactor. The remaining PAHs were analysed in a separate run at a lower split ratio with the naphthalene peak sent to waste (backflush valve on). Annotated examples of the resulting GC-C-IRMS chromatograms are provided in Figure 5.6, with the actual isotopic values for extraction repeatability presented in Figure 5.7.
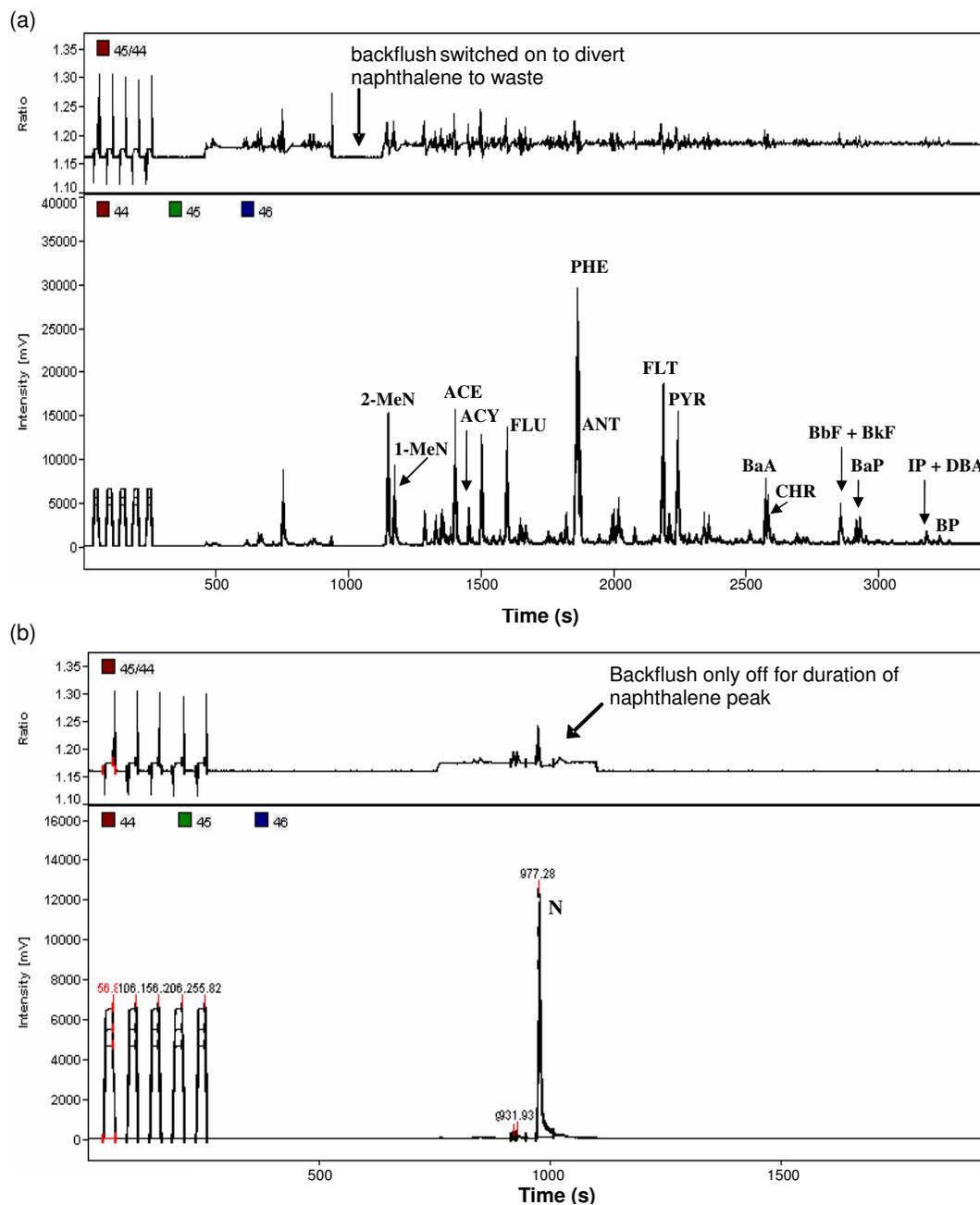
Figure 5.6: Chromatograms obtained by GC-C-IRMS analysis of fraction 2 of coal tar sample 7 using two GC methods (a) analysis of EPA PAHs at a low split ratio with naphthalene sent to waste and (b) isolation of naphthalene using a high split ratio to prevent reactor damage.
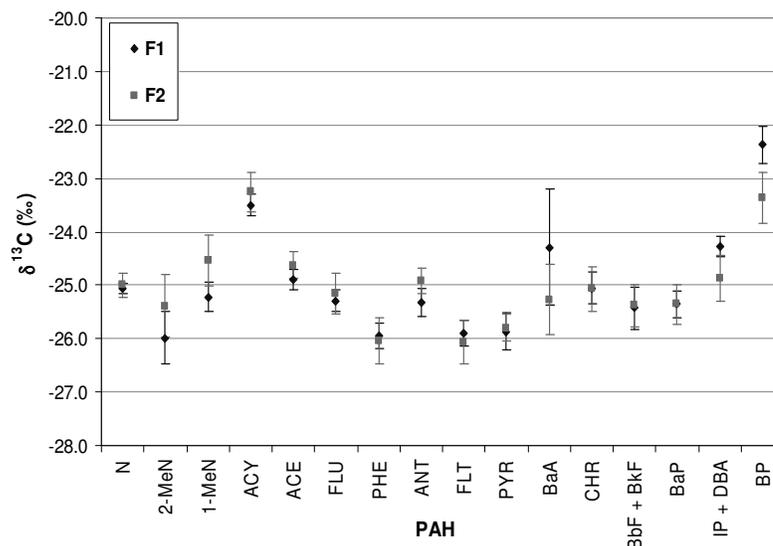
Figure 5.7: Comparison of $\delta^{13}$C values for two fractions of coal tar sample 7. Y-error bars represent a 95% CI based on six replicate fractionations with each fraction analysed in triplicate. Values are presented in Table 5.2.

As Figure 5.7 and Table 5.2 show, the carbon isotope values obtained for the 16 US EPA PAHs were consistent regardless of whether some PAHs eluted within the initial aliphatic fraction. However, the poor chromatographic performance of certain compounds, namely benzo[a]anthracene (BaA), the benzofluoranthenes (BbF + BkF) and benzo[g,h,i]perylene (BP), resulted in poor repeatability betweeen injections. This is a major problem within isotopic analysis of coal tar and was also reported by Emsbo Mattingly and Boehm [2003] despite their use of fractionation using column chromatography. The mean isotopic results obtained for fractions 1 and 2 of the six replicate extractions of coal tar sample 7 were compared using the Student's t-test using a 95% confidence limit. The results in Table 5.2 show that twelve out of the sixteen analysed PAH peaks passed the t-test, thus were deemed statistically indistinguishable [Gauchotte et al., 2011]. The four peaks which failed the t-test were the C1 methyl naphthalenes, acenaphthene and anthracene. It is possible that photolytic decomposition again caused the differences in anthracene isotopic values, however, the cause of higher variation in acenaphthene and the methyl naphthalenes is unknown. It is possible that sample evaporation has interfered with the isotopic measurements for these low molecular weight compounds, with naphthalene itself being found in such high concentration in sample 7 that it was less affected.

Table 5.2: Comparison of mean $\delta^{13}$C values for two fractions of coal tar sample 7.

| PAH | measured $\delta^{13}$C values | | | | T-test |
| | F1 | | F2 | | |
| | Mean | $\sigma^a$ | Mean | $\sigma^a$ | |
|---|---|---|---|---|---|
| N | -25.1 | 0.1 | -25.0 | 0.2 | pass |
| 2-MeN | -26.0 | 0.4 | -25.4 | 0.5 | fail |
| 1-MeN | -25.2 | 0.2 | -24.5 | 0.4 | fail |
| Acy | -23.5 | 0.2 | -23.3 | 0.3 | pass |
| Ace | -24.9 | 0.2 | -24.6 | 0.2 | fail |
| Flu | -25.3 | 0.2 | -25.2 | 0.3 | pass |
| Phe | -25.9 | 0.2 | -26.0 | 0.4 | pass |
| Ant | -25.3 | 0.2 | -24.9 | 0.2 | fail |
| Flt | -25.9 | 0.2 | -26.1 | 0.4 | pass |
| Pyr | -25.9 | 0.3 | -25.8 | 0.2 | pass |
| BaA | -24.3 | 1.0 | -25.3 | 0.6 | pass |
| Chr | -25.1 | 0.3 | -25.1 | 0.4 | pass |
| BbF + BkF | -25.4 | 0.3 | -25.4 | 0.3 | pass |
| BaP | -25.4 | 0.2 | -25.4 | 0.3 | pass |
| IP + DBA | -24.3 | 0.2 | -24.9 | 0.4 | pass |
| BP | -22.4 | --$^b$ | -23.4 | 0.4 | --$^b$ |

$^a$standard deviation
$^b$only one cell achieved levels of BP in F1 sufficient for isotope analysis, thus standard deviation could not be calculated.

The coal tar fractions were also analysed by two-dimensional gas chromatography (GCxGC) to provide a comprehensive overview of the composition of each fraction. The samples were analysed by normal phase GCxGC; a non-polar column coupled to a polar secondary column (Figure 5.8). The primary, non-polar column employed in the first dimension contained the same stationary phase as the column employed in GC-C-IRMS analyses (DB-5 equivalent phase) and the same temperature ramp program was also employed, therefore, any further separation of constituents in the second dimension (mid polar, Rtx-17 phase) shows the possible co-eluting components which may complicate the isotopic signature (refer to Chapter 6 for further details on instrumental parameters). The chromatograms obtained by analysis of ASE fractions for sample 12 are given in Figure 5.8. The chromatograms show that the aliphatic content (straight, branched and cyclic alkanes) of the sample dominates fraction 1, while the aromatic PAHs remain in fraction 2. The back plane provides the reconstituted one-dimensional GC trace, with a large unresolved hump indicating the presence much coelution within the first fraction. The chromatograms show evidence of trace amounts of components which may coelute in 1D GC, however, the quantites are very small compared to the parent PAHs thus the effect on isotopic value is assumed to be relatively low. This indicates that fast, automated ASE fractionation is acceptable for isotope studies, making the complicated, time-consuming column fractionation methods unnecessary. Thus, each sample within the coal tar library was fractionated by ASE and the second (aromatic) fractions were analysed by CSIA.
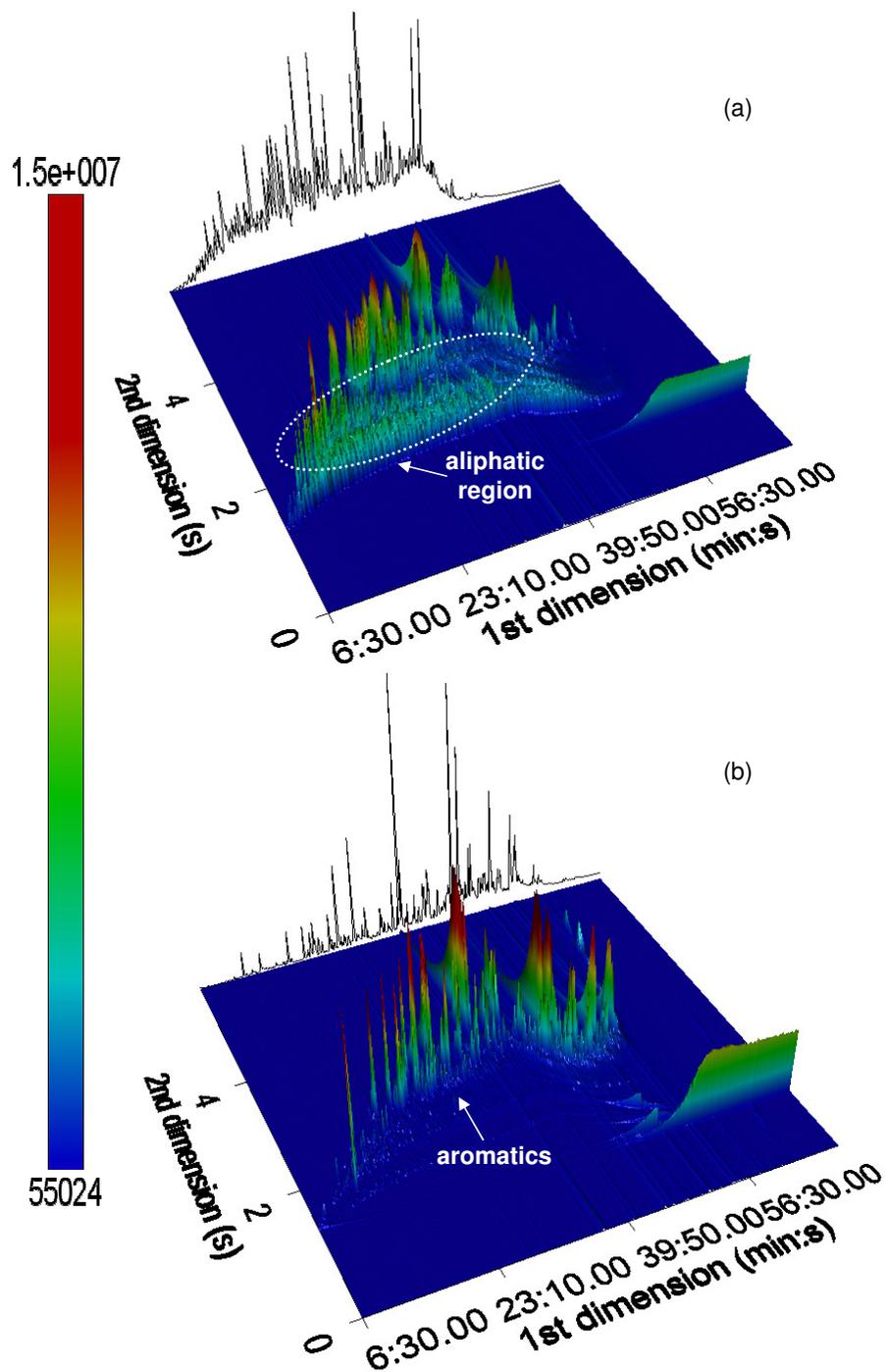
Figure 5.8: Normal phase GCxGC contour plots of (a) fraction 1 and (b) fraction 2 obtained by ASE fractionation for sample 12. The back plane shows the reconstituted one-dimensional GC trace.

## 5.3.2   Validation of fractionation method

The standard method of sample fractionation prior to CSIA is silica or alumina column chromatography. However, this method is extremely time-consuming and labor-intensive. The automated ASE procedure minimises solvent consumption to less than 100 mL per sample and two fractions are produced within 30 minutes.

Silica column fractionation was performed using coal tar sample 7 (as in the ASE repeatability tests) to allow comparison of the isotopic signatures; the results are provided in Figure 5.9. The results show that the two fractionation methods did not provide significantly different isotope values. Naphthalene showed most isotopic variation between the methods. This was expected as naphthalene was the most reactive PAH examined in the study and likely to be most affected by evaporation during sample preparation. Benzo[g,h,i]perylene also showed some isotopic variation, however, the concentrations of this analyte were generally low and provided poor chromatographic performance. The variation is most likely caused by poor instrumental precision rather than isotopic fractionation during sample preparation, thus, for the purpose of coal tar source provenancing in the remaining sections of this Chapter the isotopic values of BP will not be included. However, the full set of isotopic measurements can be found in Table B.3 of the Appendix for this chapter.
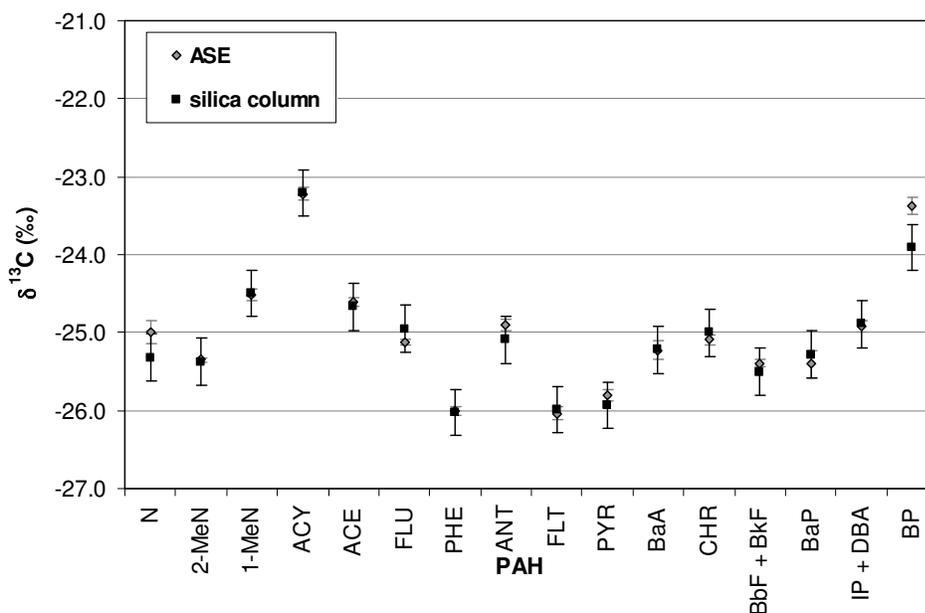
Figure 5.9: Comparison of PAH $\delta^{13}$C values obtained by ASE and silica column fractionation of coal tars (using aromatic fraction 2 of coal tar sample 7). Y-error bars represent 0.3 ‰ for the silica column method and a 95 % CI based on triplicate injections for the ASE method.

### 5.3.3   Comparison of coal tar isotopic signatures

It was decided, after evaluation of results from the described repeatability tests, that the automated fractionation method was acceptable for the purpose of CSIA of coal tars. Therefore, the carbon isotope values of the aromatic fraction (F2) of each coal tar from the library was measured and evaluated to determine whether trends in carbon isotope value could be established, both across single sites and between different FMGP sites. The results are summarised in Figure 5.10, but the raw data is also supplied in Appendix B. As Figure 5.10 shows, the isotopic signatures varied between the coal tar samples but were generally confined to the narrow isotopic range of -24 to -28 ‰. The most obvious differences in isotopic signature are apparent in samples 2, 3 and 15, which are isotopically light compared to the remaining tars. The isotope values obtained for PAHs in samples 2, 3 and 15 are closer to $\delta^{13}$C values generally associated with petrogenic origin (c.a. -28 to -30 ‰).
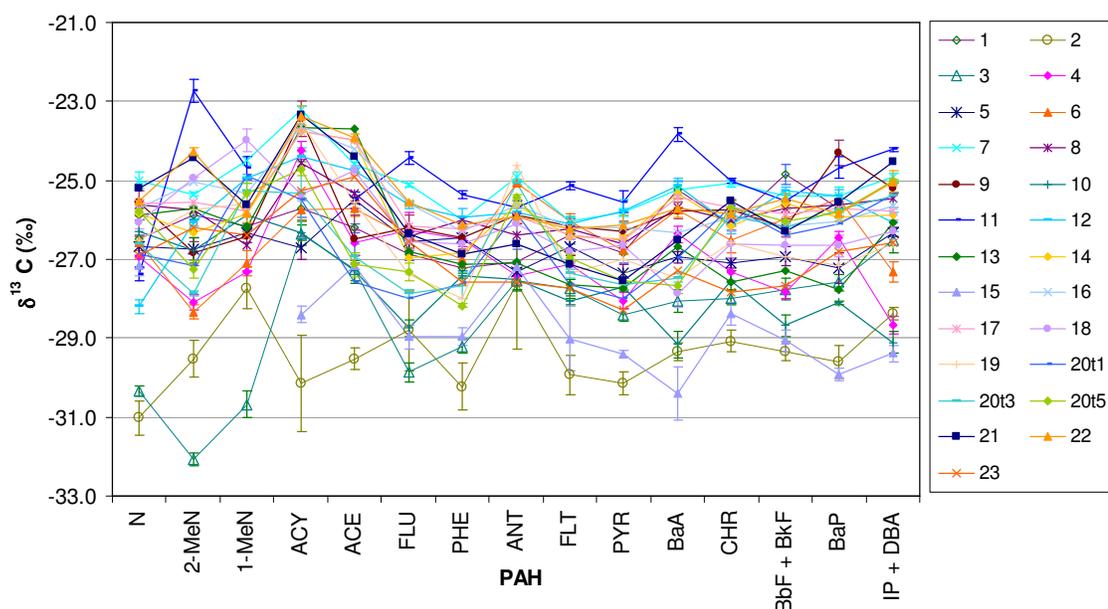
Figure 5.10: Carbon isotopic values of the dominant PAHs in all 25 coal tar samples. All values were measured using aromatic fraction 2 of the extracted tars.

A number of plots were constructed to compare the isotope ratios of the most well-resolved PAH compounds with good peak shape; as presented in Figure 5.11. The majority of coal tars cluster within a large group (as indicated by the arbitrary boundaries) however, samples 2, 3, 15 and 11 are generally located outside of the main cluster. This demonstrates a fast method of comparing the isotopic compositions to allow any differences to be explored more thoroughly. Figure 5.11 demonstrates a linear relationship between the well-resolved PAHs, in particular fluoranthene and pyrene.

Samples 1-6 (from site S1) are highlighted in Figure 5.11(a) and are all located along the linear trendline. It is possible that the trendline represents the degree of mixing with a petrogenic source. Sample 2, which has the lightest isotopic values in the coal tar library, is positioned on the left of the graph, while creosote is at the opposite end of the trendline. This form of isotopic evaluation could prove useful for tracking different coal tar DNAPL plumes and plume mixing at FMGP sites.
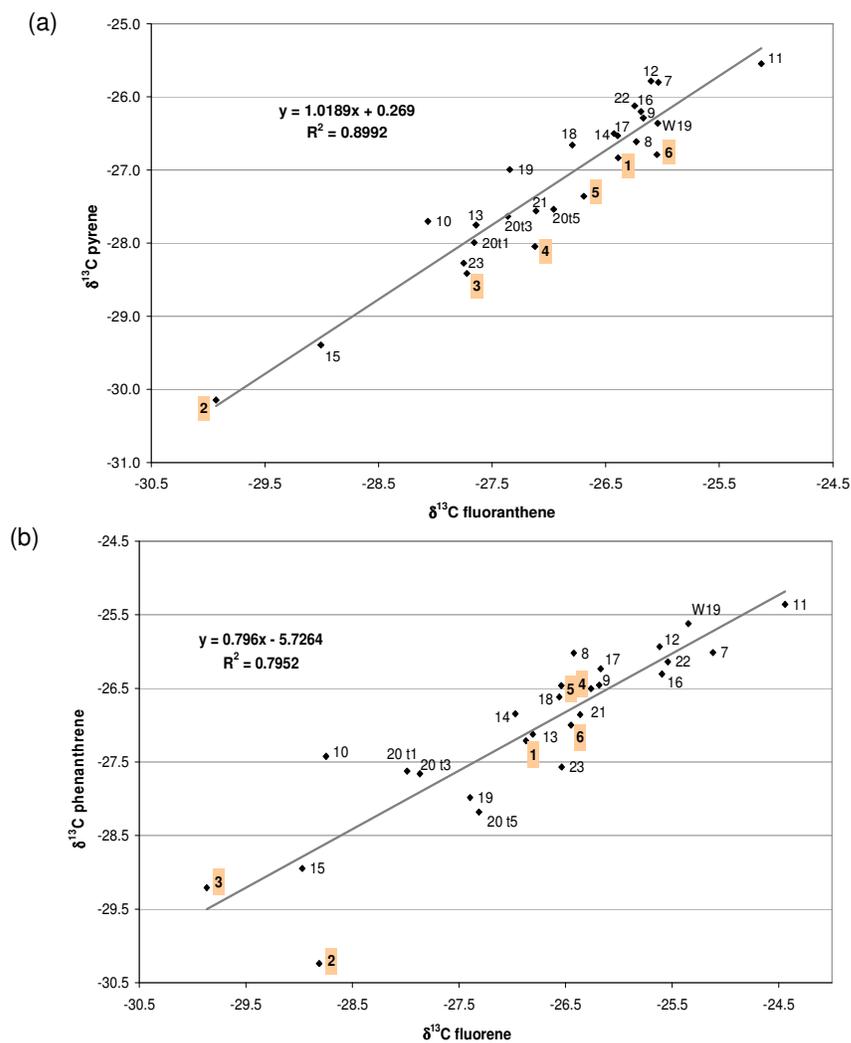
Figure 5.11: Comparison of carbon isotope compositions of selected well-resolved PAHs (a) fluoranthene and pyrene and (b) fluorene and phenanthrene. All values are presented in units of per mille (‰).

**Trends across single sites**

Coal tar samples 1-6 were obtained from different locations within the same site (S1), thus were expected to have similar signatures. However, the isotopic values showed much variation (Figure 5.12). In particular, coal tars 2 and 3 gave lower isotope values, with many PAHs exhibiting $\delta^{13}$C values less than -29 ‰. Such values are within the range normally associated with petroleum products [McRae et al., 1996]. Sample 3 appears to be a mixture of two sources, the majority of lighter PAHs have isotope values characteristic of a petrogenic source, while the higher molecular weight PAHs appear to be of a pyrogenic (coal tar) character. Sample 3 was obtained from a borehole within an old tar tank, therefore, it was expected to be associated with vertical coal retort processes on the site. However, the site is known to have also used petroleum products, such as primary flash distillate, for CWG processes which may have mixed with the coal tar in the subsurface environment. On the other hand, sample 2 was obtained from a borehole not strongly associated with the vertical retorts, and as such has a greater chance of having been affected by multiple sources. Consequently, CSIA of coal tar PAHs from many samples across a single site could help to uncover multiple contaminant plumes and aid in identification of source.
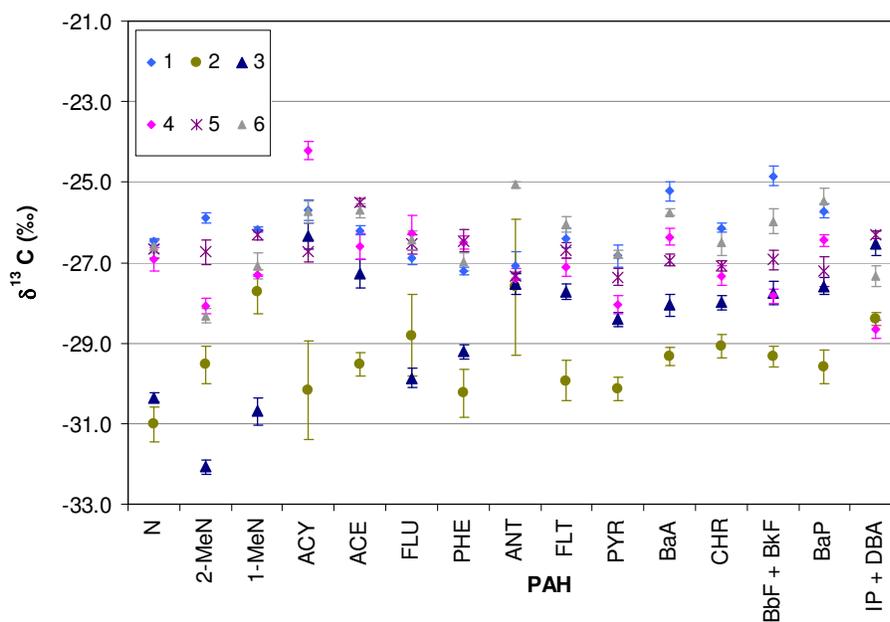


Figure 5.12: Comparison of carbon isotope signatures of PAHs in coal tar samples 1-6 from site S1. Y-error bars represent a 95% CI based on triplicate injections.

Coal tar samples 20t1, 20t3 and 20t5 were all obtained from site S13 at the same sampling location, thus were expected to have similar carbon isotope signatures (Figure 5.13). As anticipated the isotopic signatures are strongly correlated, with only a slight variation amongst the isotope values for acenaphthylene (ACY) which was most likely due to poor chromatographic performance for low intensity peaks rather than an actual variation in isotopic signature of the tar.
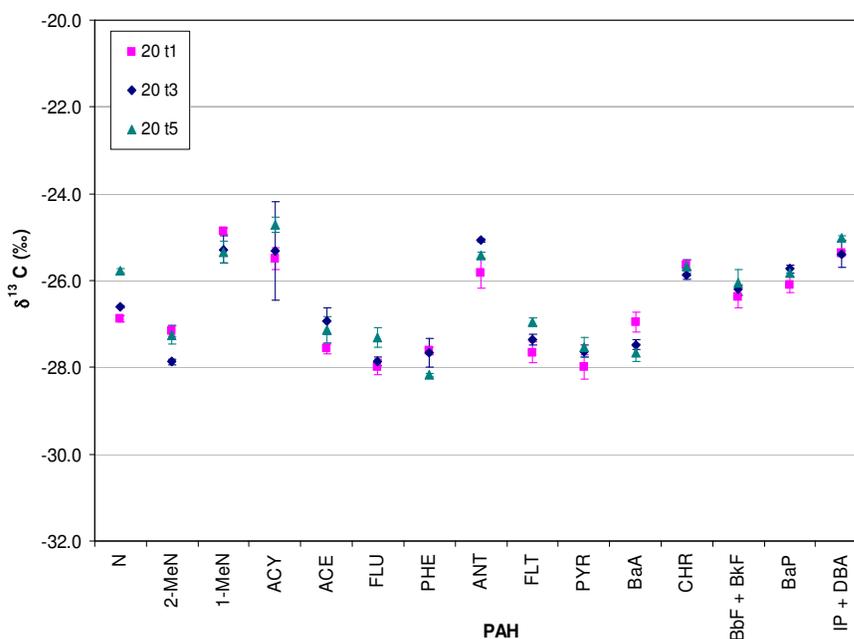


Figure 5.13: Comparison of carbon isotope signatures of PAHs in coal tar samples from site S13. Y-error bars represent a 95% CI based on triplicate injections.

Samples 14 and 17 which were obtained from the same FMGP site (S9) displayed interesting isotope signatures (Figure 5.14). Certain PAH isotope values, such as for acenaphthylene (ACY) and acenaphthene (ACE), are closely correlated for the samples while others are distinctly different. Carbon isotope ratios for acenaphthene and acenaphthylene frequently appear less negative than the other monitored PAHs. It is possible that these compounds are produced as primary devolatilisation products, while other PAHs are formed by secondary condensation reactions, such as ring closure or ring fusion, which are more likely to be affected by isotopic fractionation [McRae et al., 1999]. Formation of C-C bonds would be affected by the kinetic isotope effect, as $^{12}C$ would be preferentially incorporated into the bonds, thus giving $^{12}C$-enriched isotope values for the resulting PAHs. Therefore, it is possible that ACY and ACE were produced in both samples 14 and 17 as primary devolatilisation products, meaning that if the same coal supply was used across the site
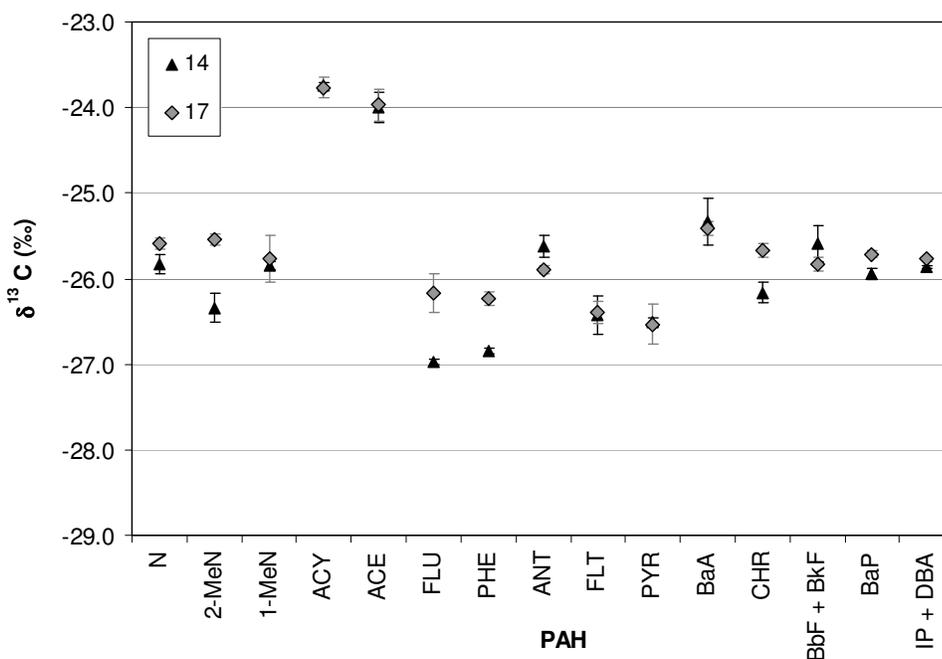
Figure 5.14: Comparison of carbon isotope signatures of PAHs in coal tar samples from site S9. Y-error bars represent a 95% CI based on triplicate injections.

(which is highly likely) the values should be similar. However, the sampling locations of tars 14 and 17 suggest their formation by different manufacturing processes, CWG and horizontal retorts respectively. Consequently, differences in manufacturing process parameters could have resulted in different levels of isotopic fractionation. Regardless of this fact the variation in isotopic signature is not so distinct as to allow unambiguous discrimination between different manufacturing processes, especially as weathering processes may also have affected the values.

**Trends by manufacturing process**

The carbon isotope values are very similar across all FMGP sites investigated, generally within the narrow range of -24 to -28 ‰. CSIA of coal tar PAHs did not appear to provide adequate source-specific identification in this study. Therefore, the coal tar samples were grouped by historic manufacturing process and average carbon isotope values were calculated for each group to discover if trends between manufacturing process occurred. Figure 5.15 shows the trend in $\delta^{13}$C values for the six manufacturing processes investigated in this study.

The mean isotope values presented in Figure 5.15 show little variation between manufacturing processes. Figure 5.15 illustrates the differences caused by manufacturing process within individual PAHs which were well resolved in the coal tar chromatogram. One of the main distinctions in carbon isotope values is between the creosote sample and other coal tars. Creosote is a distilled fraction of coal tar, thus it has greater opportunity for isotopic fractionation. Surprisingly, the isotopic composition of CWG tars was not significantly different from other coal tars. At CWG sites, large quantities of oil were used to enrich the gas stream. According to McRae et al. [1996], petroleum oils generally exhibit lower $\delta^{13}$C values (in the range of -28 to -31 ‰) than parent coals (c.a. -25 ‰). It was anticipated that this would allow differentiation of CWG tars, however, this was not the case. It is possible that weathering processes resulted in isotopic fractionation and alteration of the isotopic signature. The most likely explanation is due to the fact that light oils were generally used for carburetion processes at FMGPs in the British Isles, rather than the heavy oils which were used in the United States. As previously mentioned, light oils are unlikely to have contributed significant amounts of tar at CWG sites compared to the quantities of tar which heavy oils were capable of producing. Therefore, CWG tars from British sites are more likely to represent the coal/coke used within the initial stage of the CWG process prior to introduction of the oil spray. This also explains the results achieved by Saber et al. [2006] who noted the distinction between CWG and coal carbonisation tars, as these were obtained from FMGP sites in the US which are likely to have utilised heavy oils for gas enrichment.
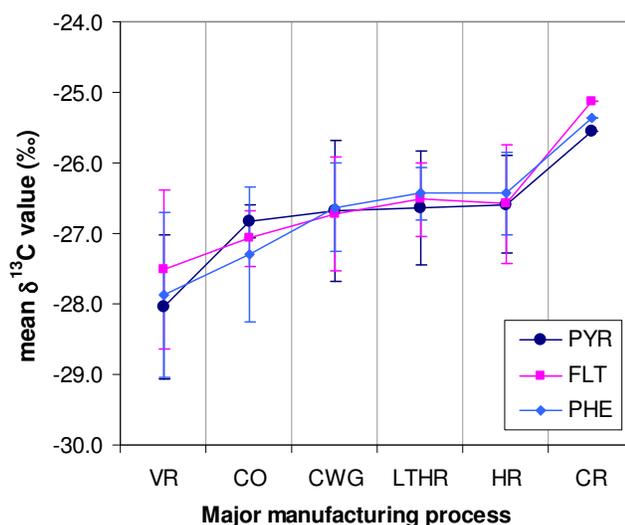
Figure 5.15: Relationship between carbon isotopes and historical manufacturing process for specific well-resolved PAHs [where CO=coke oven, CR=creosote, CWG=carburetted water gas, HR=horizontal retorts, VR= vertical retorts and LTHR=low temperature horizontal retorts]. Y-error bars represent standard deviation based on triplicate injections.

All coal tars investigated in this study were obtained from FMGP sites within the British Isles, therefore, it is safe to assume that mainly British coals were used for gas production. MGPs are likely to have used local coal suppliers to limit expense and although the coal tars in this study were obtained from sites across the whole of Great Britain, it is unlikely that regional variation in isotope signature of parent coal would be dramatic enough to yield differences in isotopic composition of the resulting coal tar. The similarity of raw feedstocks may be the cause of poor source discrimination. It is possible that a larger dataset of tars including worldwide sources may allow a trend in isotope signature with coal region to be discovered. However, the expansion of the dataset in this manner would be most useful to discover differences in isotopic signature due to the use of alternative feedstocks (such as oil or wood) for gas production when coal was in limited supply; CSIA may provide differentiation of such sites, however this was not the case for the coal tar library investigated in this study. It is thought that the similar nature of raw materials (i.e. British coals) produced similar isotopic signatures in the resulting tar, thus preventing effective source apportionment.

Coal tar samples 18 and 19 were obtained from coke ovens at a present-day steel works. The tars were produced on the day of sampling, thus represent the freshest samples within the coal tar library and have not been exposed to weathering in the subsurface environment as with all other

tar samples. The comparison of samples 18 and 19 show similar isotopic signatures despite the different oven types and temperatures used in their manufacture. The carbon isotope values for these samples were also isotopically lighter than from other coal tars, with average PAH composition of -26.4 ‰. However, the coal used to fuel the coke ovens was sourced from America and Australia, thus it is expected that their isotopic values would vary from those of British coals. British coals have previously been found to have average carbon isotope values of -24 to -25 ‰ [McRae et al., 1999]. This is promising for source provenancing of more recent coal-derived contamination where multiple types of coal from many regions in the world may have been used, e.g. at present day steel or coking works.
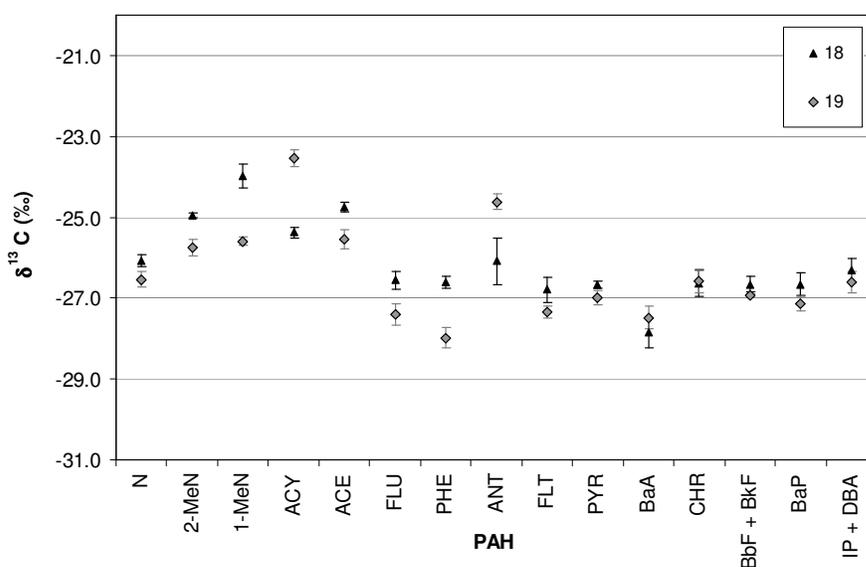


Figure 5.16: Comparison of fresh coke oven tars and weathered sample of coal tar 19 (labelled W19).

As previously mentioned, the isotopic composition of coal tars has not been extensively reported in literature. The results demonstrated in this study show that it is difficult to allocate source based solely on the PAH isotopic composition. The technique shows potential for the development of mixing models at FMGPs to highlight the different PAH sources (both petrogenic and pyrogenic) present across a single site. However, the addition of isotopic measurements for further elements may also increase the discriminatory power of the technique. For example, as the major elements within coal tar are carbon and hydrogen, the most logical choice would be to measure the $\delta$D values of the PAH compounds.

## 5.4    Conclusion

In this Chapter, compound specific carbon isotope analysis was performed on the PAH fractions of all 25 coal tar samples. The complex nature of coal tars requires fractionation of extracts prior to analysis, thus an automated sample fractionation method was developed using accelerated solvent extraction (ASE) in a bid to save time and reduce solvent consumption. Comparison of the ASE method with traditional column chromatography demonstrated that fractionation method did not significantly alter the carbon isotope values.

The narrow range of $\delta^{13}C$ values and poor instrumental performance (such as poor peak shape and coeluting compounds) for coal tar PAHs makes source apportionment of coal tars via this technique extremely challenging. It was not possible to attribute coal tars to certain sites based solely on carbon isotope values. Furthermore, the results did not demonstrate significant variation of $\delta^{13}C$ value with historical manufacturing processes due to the similarity between raw materials used in the process.

The potential of CSIA for coal tar analysis was evident in the analysis of multiple tar samples from various locations across a single FMGP site. The carbon isotope values showed clear differences between the tars, indicating the possibility of multiple contamination sources and the mixing of sources. Furthermore, the carbon isotope range of coal tar demonstrated by this study is sufficiently different from reported values for petroleum products to allow differentiation of multiple sources if samples are too degraded to provide accurate chemical fingerprints for pattern recognition via GC-FID or GC-MS analyses. The development of a larger, global database of coal tars may uncover further trends in isotopic signature, however, this was outside the scope of this thesis. Differences in the isotopic signature of coal tar due to differing raw materials may be discovered. For example, when heavy oils were used as a gasworks feedstock, due to limited coal supplies, isotopically light values (associated with petroleum products) would be expected.

In the following Chapter, an ultra resolution method of coal tar chemical fingerprinting with the capacity to differentiate tar source is discussed. The method employs another advanced analytical technique with great potential for environmental forensic applications, known as comprehensive two-dimensional gas chromatography.

# Chapter 6

# Reversed Phase Two-Dimensional Gas Chromatography

## 6.1 Introduction

The analysis of complex mixtures is a major challenge within many scientific fields. Conventional gas chromatographic techniques are capable of high peak capacity, yet they struggle to provide separation of individual constituents present in complex mixtures such as crude oil or coal tar.

Comprehensive two-dimensional gas chromatography (GCxGC) is a high resolution separation technique, developed with the intention of overcoming limitations associated with conventional GC techniques. The coupling of two columns with different selectivity allows for a two-dimensional orthogonal separation of mixtures, across a retention plane rather than along a retention line [Phillips and Beens, 1999, Marriott and Shellie, 2002, Adahchour et al., 2006b,c, Dalluge et al., 2003]. This allows an order of magnitude more compounds to be separated by GCxGC than when using conventional gas chromatographic instrumentation [Bertsch, 1999]. A full detailed description of the GCxGC technology is provided in Chapter 3.

The two capillary columns are connected by a modulator which traps and focusses first column eluents prior to re-injection onto the secondary column. This process occurs approximately every

135
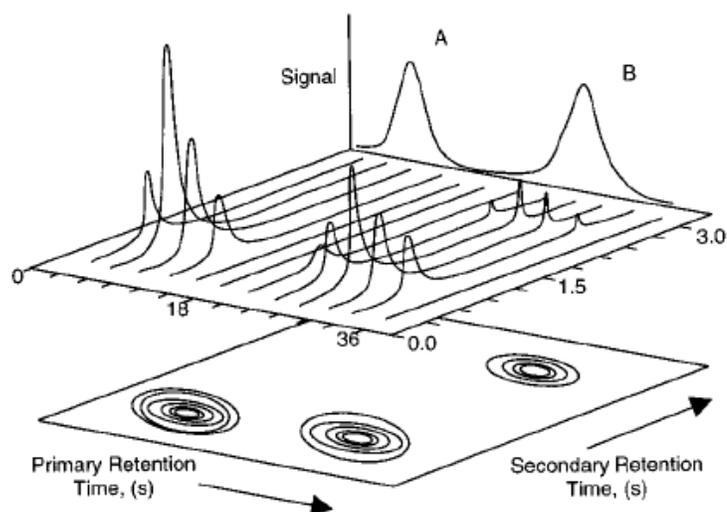
Figure 6.1: A schematic of GCxGC chromatogram visualisation. The backplane represents the one-dimensional GC trace, which is further separated in the second dimension, while the lower section represents the final contour plot (copied from Gaines et al., 1999).

4-6 seconds, depending on the modulation period set by the user. Therefore, the output of GCxGC combines a series of second dimension separations stacked side-by-side to produce a two-dimensional retention time plane [Gaines et al., 1999]. A colour gradient is then added to indicate signal intensity, allowing the retention plane to be viewed as a 3D surface plot. An alternative representation is the colour contour plot, which provides a birds-eye-view looking down on the peaks. An illustration depicting the formation of a GCxGC chromatogram is given in Figure 6.1.

Generally, a long, wide bore (0.25-0.32 mm internal diameter, id), non-polar capillary column is used in the first separation, whereas a short, narrow bore (0.1-0.2 mm id), polar column is installed for the second separation; this is deemed normal phase. The short, narrow secondary column allows for fast separations in the second dimension. As GCxGC is a relatively new technique, operators tend to rely on conventional wisdom when optimising certain parameters. For example, very narrow secondary columns (0.1 mm id) are generally chosen, even though it has been shown that they can become overloaded easily [Harynuk et al., 2005]. It is up to the operator to optimise the system for each specific application. The optimisation of a GCxGC system for the analysis of coal tars will be discussed in this Chapter.

The use of normal phase separation is standard procedure in GCxGC analyses, as this is typically

136
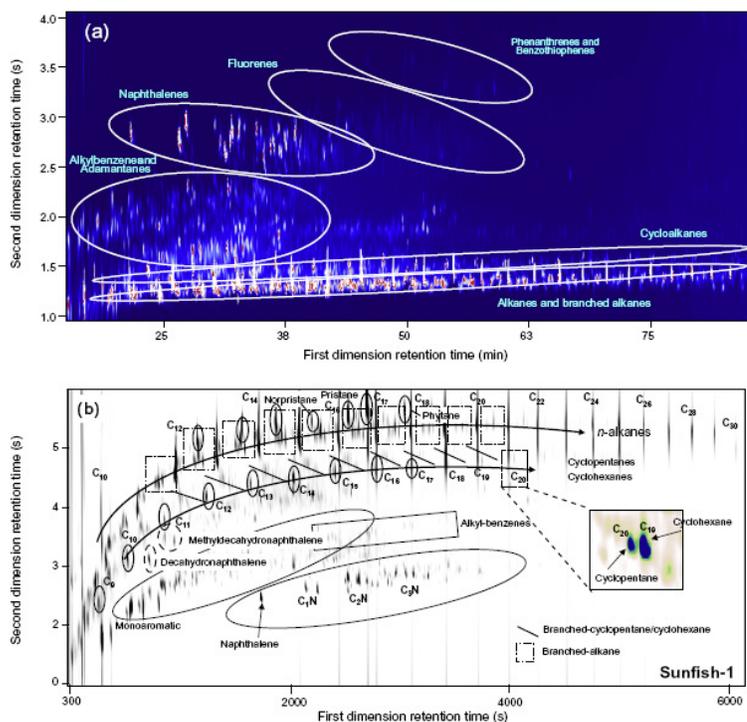
Figure 6.2: Contour plots obtained by GCxGC analyses of crude oil using (a) normal phase and (b) reversed phase [copied from Tran et al., 2010].

the column set supplied with commercial instruments. However, inverting the column phases has been shown to provide better group-type separation in certain cases [Van Der Westhuizen et al., 2008, Tran et al., 2006, 2010]. The use of a polar primary column and non-polar secondary column is known as reversed phase (or reversed polarity) GCxGC [Adahchour et al., 2006b]. Tran et al. [2010] analysed a range of crude oil samples in varying degradation states by reversed phase GCxGC. The study showed that normal phase gave poor use of the chromatographic space, thus a reversed phase column set was employed. Reversing the column polarity improved orthogonality and the resulting chromatographic separation. In particular, the aliphatic region which had previously been closely packed due to fast elution from the polar secondary column, was well resolved by a reversed phase column set. The difference in orthogonality of the two methods is illustrated in Figure 6.2.

### 6.1.1   Application in Environmental Forensics

GCxGC technology has already been applied within a vast range of scientific disciplines; including varied applications such as doping control [Mitrevski et al., 2010], arson investigation [Frysinger

and Gaines, 2002] and flavour/fragrance technology [Dalluge et al., 2003]. However, the technique has been shown to be especially useful for environmental analyses [Frysinger et al., 2002, Skoczyska et al., 2008]. The main advantage of GCxGC, when applied to environmental forensics, is minimisation or elimination of fractionation processes prior to analysis Muhlen et al. [2006], Panic and Gorecki [2006]. A complex sample can be injected as a single extract without involving time-consuming fractionation processes. This gives fast screening of the entire sample, allowing many classes of organic contaminants to be monitored simultaneously. GCxGC FID has been shown to be useful as a screening technique for oil contamination as it provides good separation at a relatively low cost [Ong et al., 2003]. However, coupling to a mass analyser allows an additional level of information on the chemical structure of individual constituents to be obtained [Ong et al., 2003]. Currently, a time-of-flight mass spectrometer (TOFMS) is the only mass analyser capable of achieving spectral acquisition rates (100-200 Hz) sufficient for the fast secondary separations in GCxGC analyses [Phillips and Beens, 1999].

Two-dimensional gas chromatography is well-suited to the analysis of crude oil and other petroleum products, as these represent some of the most multi-class samples available [Muhlen et al., 2006]. Gaines et al. [1999] found that the structured nature of GCxGC chromatograms (chemical families elute in grouped bands) is extremely beneficial for comparing the composition of oil samples. Oil spill samples contain similar chemical classes to those found in coal tars, such as n-alkanes, iso-alkanes, alkenes, cyclic alkanes, aromatics and heterocycles, thus, GCxGC is an interesting prospect for coal tar analyses. Regardless of such advances in analytical technology, there have been few reports on coal tar composition in the literature. Most recently, Brown et al. [2006] evaluated the composition of coal tars from ten different FMGP sites in the US, indicating major differences in PAH composition between sites. However, this study utilized GC-MS analyses, after conventional sample fractionation, so the chemical information obtained on the coal tar was limited by resolution power and restricted mainly to the 16 priority pollutant PAHs.

To the author's knowledge, this work represents the first application of GCxGC for source apportionment of coal tars [McGregor et al., 2011a,b]. However, Vasilieva et al. [2011] have recently used GCxGC to monitor biodegradation in aromatic tar oils, such as anthracene oil. The study investigates the enhanced separation of GCxGC to monitor all contaminants in a single analytical run. However, the column set was restricted to a basic, normal phase arrangement and the detector in use was a quadrupole mass spectrometer (GCxGC/qMS). Quadrupole MS is not capable of reaching

the fast acquisition rates required to measure the secondary column eluent. The combination of poor column set and slow spectral acquisition resulted in poor peak shape in the second dimension; this is evident in the contour plot shown in Figure 6.3.
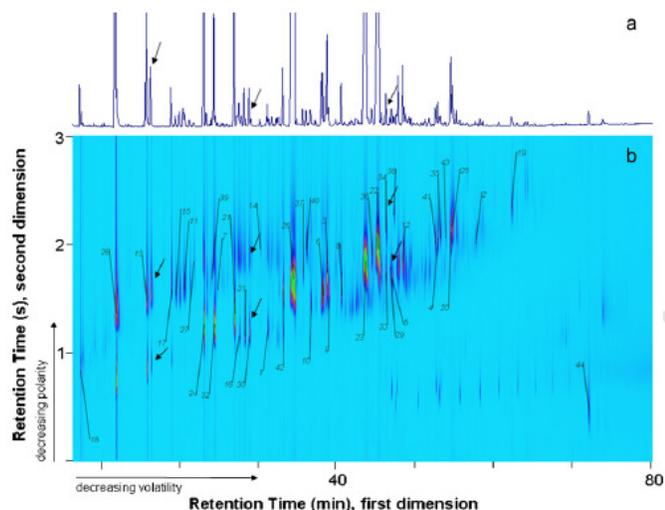


Figure 6.3: GCxGC/qMS chromatogram of tar oil contaminated soil. Arrows show where co-eluting peaks in the one-dimensional chromatogram (a) become resolved into separate peaks in two-dimensional separation (b) (copied from Vasilieva et al., 2011).

Another recent publication by Machado et al. [2011] focussed solely on the identification of sulfur-containing compounds in coal tar. The study aimed to resolve problems associated with the co-elution of sulfur heterocycles and PAHs when using conventional GC techniques. However, the sulfur heterocycles were not entirely resolved. The study employed a basic normal phase column set which is likely to have restricted separation of the components. Furthermore, the paper was limited to the examination of sulfur compounds (which make up only a small portion of the total composition) in coal tars produced by lab-scale pyrolysis studies, rather than real FMGP samples which would have been exposed to weathering.

This Chapter aims to resolve the issues associated with the analysis of complex coal tars through the use of reversed phase GCxGC TOFMS. GCxGC TOFMS can provide accurate and precise chemical fingerprinting of coal tars by enhanced separation of their individual components. Chemical fingerprinting of environmental samples by conventional GC techniques is often described as high resolution separation. In this work, an improved method of chemical fingerprinting, deemed 'ultra resolution', is developed. The process gathers more chemical information per sample than traditional tiered approaches and has the additional benefits of using an automated one-step extraction

and analyses using a single instrument.

### 6.1.2    Aim and Objectives

The overall aim of this Chapter was to explore the use of two-dimensional gas chromatography as an ultra resolution method of chemical fingerprinting for coal tars. The objectives set to reach this aim were:

- The optimisation of extraction procedures for coal tar samples.

- The comparison of GCxGC TOFMS with conventional gas chromatography (as applied to coal tar analysis).

- The optimisation of instrumental parameters for optimal separation of coal tar components.

## 6.2    Materials and Methods

### 6.2.1    Sample Preparation

Extraction was performed using an ASE 350 accelerated solvent extraction system (Dionex, Camberley, UK) equipped with 10 mL stainless steel extraction cells. The high separation capability of GCxGC TOFMS eliminates the requirement for sample fractionation, thus a single extraction using hexane (including in-cell cleanup by silica gel) was performed.

Anhydrous sodium sulfate, silica gel 60 (both from Sigma Aldrich) and diatomaceous earth (Dionex, Camberley, UK) were activated for 4 hours at 400-450 ˚C prior to use. Silica gel 60 was then deactivated by 10 % (w/w) using deionised water.

A dry, homogeneous mix of each sample was prepared by grinding the coal tar (approximately 0.5 g) with sodium sulfate ($Na_2SO_4$) and diatomaceous earth (D.E.) in a 1:1:1 ratio. $Na_2SO_4$ removes any residual water content, while grinding with D.E. results in fine powder (rather than a tar) which can be transferred quantitatively to the extraction cells. The US EPA recommend the use of Method 3580, Waste Dilution [US EPA,1997] for the preparation of DNAPL samples, however,

for this approach a simple solvent dilution is employed and requires subsequent cleanup of extracts to remove unwanted particulates. To ensure accurate quantification was achieved, each coal tar mixture was spiked with 600 $\mu$L of the stock surrogate solution (2000 $\mu g/mL$ of D8-naphthalene, D10-fluorene, D10-fluoranthene and D12-chrysene in dichloromethane) prior to grinding with D.E. and $Na_2SO_4$. This allows any loss of target analytes to be monitored from the start of sample preparation. Furthermore, storage of the sample in this form allows the loss of target analytes over time to be monitored.

ASE cells were lined with 2 filter papers (to trap unwanted particulate matter) and packed with 3 g silica gel 60 (10% deactivated w/w). Approximately 0.5 g of the ground coal tar/surrogate mixture was added to each extraction cell and the remaining cell volume was packed with D.E. Hexane was used as the extracting solvent for all extractions. ASE was performed at 150 °C and 10 MPa, using one dynamic (7 min) and two static extractions (5 min each). A flush volume of 150 % and purge time of 60 s were used. The extracts were concentrated to 1 mL using a Büchi Syncore® Analyst (Oldham, UK). The extracts were then made up to an exact volume (between 2-10 mL) using hexane. A 1 mL aliquot was then transferred to an autosampler vial and spiked with 75 $\mu$L of internal standard (D10-phenanthrene) prior to analysis.

## 6.2.2   GC-MS analyses

A Thermo Scientific (Hertfordshire, U.K.) Trace Ultra GC fitted with a DSQII mass spectrometer and Triplus autosampler was used for all GC-MS analyses. The column was a J&W Scientific DB-5 fused silica capillary column (30 m x 0.25 mm id x 0.25 $\mu$m film thickness). All injections were of one microlitre and were carried out using a split ratio of 1:50 and injection port temperature of 230 °C. Helium was used as the carrier gas, with a flow rate of 1 mL/min. All standards and extracts were analysed with the oven temperature programmed at 10 °C /min from 55 °C (2 min isotherm) to 110 °C, 3 °C/min to 210 °C, then at 8 °C/min to 310 °C (15 min isotherm). The ion source and transfer line were maintained at 200 °C and 320 °C respectively. The electron ionization voltage was set at 70 eV for all analyses.

### 6.2.3   GCxGC TOFMS analyses

All GCxGC TOFMS analyses were performed using a Leco (St. Joseph, Michigan) time-of-flight mass spectrometer, model Pegasus 4D, connected to an Agilent 7890A gas chromatograph equipped with a Leco thermal modulator. The TOF ion source was fixed at 200 °C and masses between 45 and 500u were scanned at a rate of 200 spectra per second. The detector voltage was set at 1700 V and the applied electron ionization voltage was set at 70 eV. All results were collected and viewed using ChromaTOF® software (Leco, version 4.22).

All standards and extracts were analysed with the primary oven temperature programmed at 10 °C /min from 55 °C (2 min isotherm) to 110 °C, 3 °C/min to 210 °C, then at 8 °C/min to 310 °C (15 min isotherm). The secondary oven was programmed at 10 °C /min from 75 °C (2 min isotherm) to 130 °C, 3 °C/min to 230 °C, then at 8 °C/min to 330 °C (15 min isotherm). This corresponds to a 20 °C offset between the primary and secondary ovens. The modulator temperature was also maintained at a 20 °C offset relative to the primary oven and the modulation period was 6 seconds with a 1.3 second hot pulse time. The transfer line was maintained at a temperature of 340 °C. The injection port temperature was set to 250 °C using a split ratio of 1:50. One microlitre of sample was injected for each run using an MPS2 twister autosampler (Gerstel). Helium was used as the carrier gas, with a flow rate of 1 mL/min.

Primary and secondary columns were connected via a Thames Restek Press-tight® glass connector. Numerous column sets were evaluated during the optimisation of coal tar separation; specific column parameters are listed in Table 6.1 and will be discussed in the following sections.

## 6.3   Results and Discussion

### 6.3.1   Optimisation of Extraction

The enhanced separation capacity of GCxGC negates the requirement for rigorous sample fractionation steps. Complex mixtures can be analysed as a single extract. Accelerated solvent extraction (ASE) is an automated method of sample extraction which uses elevated temperature and pressure to help break analyte to matrix bonds. Previous literature on ASE has shown that the technique

can provide exhaustive extraction of PAHs from soil samples with high efficiency and no carry over [Ong et al., 2003]. The study by Ong et al. [2003] also evaluated the use of simultaneous extraction and cleanup, concluding that it achieved equal (if not higher) yields of PAHs with respect to conventional cleanup methods using silica/alumina column chromatography. For this reason, it was decided to apply ASE offline GCxGC TOFMS for the study of coal tar composition.

The aim of the ASE procedure was to provide a single coal tar extract, containing all organic chemical classes in a fast, efficient manner. An ASE cycle can be completed in approximately 25 minutes, while conventional methods, such as Soxhlet extraction may take up to 48 hours for a single sample [Eskilsson and Bjorklund, 2000]. Time reduction is not the only benefit of ASE, solvent consumption for exhaustive extraction is also significantly improved. The ASE method requires only 35 mL of solvent per sample, including rinse volume, which is a fraction of the solvent volume required by traditional extraction methods. For example, Soxhlet extraction requires between 100-500 mL of solvent per extraction [Eskilsson and Bjorklund, 2000]. The ASE procedure also employed simultaneous extraction and clean-up, further reducing the total analysis time and solvent consumption.

Initial ASE parameters were based on the US EPA SW-846 Method 3545A for pressurised liquid extractions [USEPA, 2007] and adapted to give optimal results for coal tar GCxGC. It was found that hexane was suitable as the extracting solvent. Extraction cells re-extracted using dichloromethane, however, only the internal standard peak was evident in the gas chromatogram of the extracts (as shown in Figure 6.4). This demonstrates that the method provided exhaustive extraction. The ease of dissolution of coal tars in hexane avoids the need for harmful chlorinated solvents, such as dichloromethane.

Four deuterated PAHs (D8-naphthalene, D10-fluorene, D10-fluoranthene and D12-chrysene) were chosen to monitor extraction efficiency as they span a range of molecular masses, from 136 g/mol to 240 g/mol, allowing the effect of molecular weight on recovery to be monitored. Calibration of the GCxGC instrument was performed to allow quantification of each of the 16 EPA PAHs and surrogates within the final extracts; an example set of calibration curves are provided in Appendix A. A new set of calibration curves were prepared for each batch of analytical samples to account for any instrumental variability.

Repeatability of the extraction method was evaluated in two ways; repeat extraction of PAH-spiked
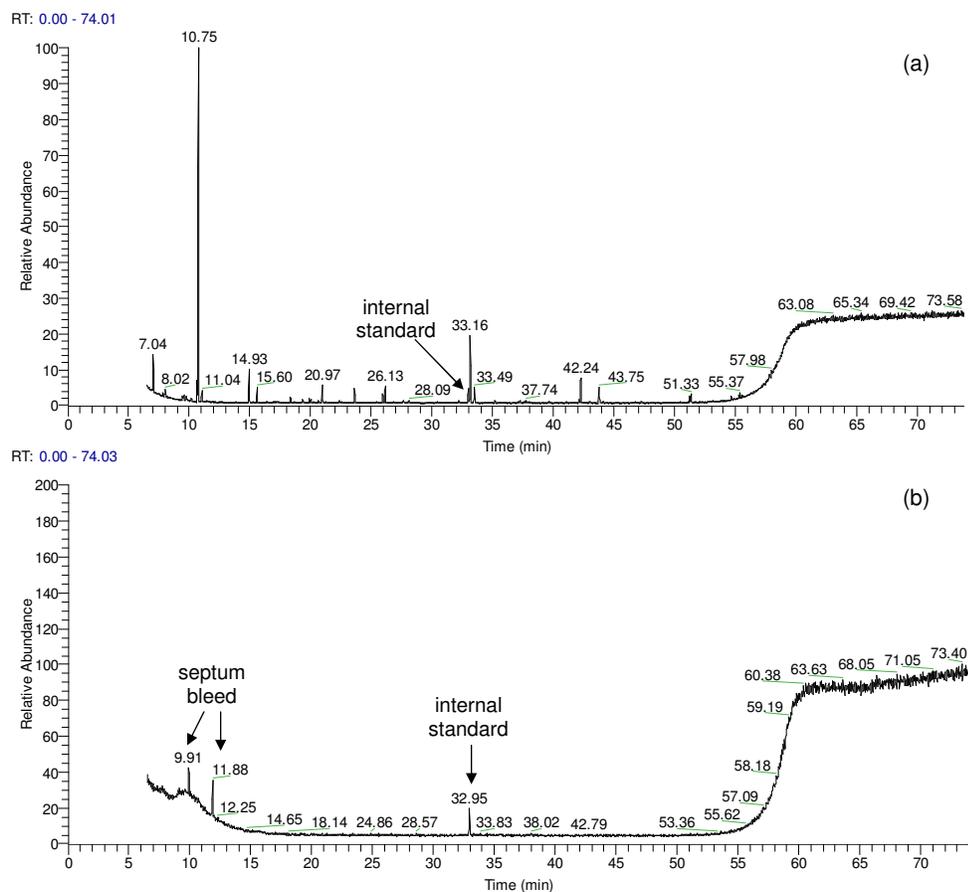
Figure 6.4: GC-MS chromatograms showing (a) the full coal tar extract in hexane and (b) the results obtained through re-extraction of the ASE cell with dichloromethane for coal tar sample 7.

cells and analysis of replicate cells containing a coal tar sample. For the PAH-spiked replicates, six cells were prepared by grinding diatomaceous earth and sodium sulfate with a 300 $\mu$g spike of each of the 16 EPA PAHs as it is very difficult to replicate a blank coal tar matrix. Prior to analysis the extracts were concentrated to 1mL and rediluted to 2 mL, thus if 100% extraction efficiency was achieved, the final analysed extract should theoretically contain each PAH at a concentration of 150 $\mu$g/mL. The extraction efficiency for each PAH was calculated as a percentage based on these theoretical values. The mean extraction efficiencies for all PAHs ranged from 80-110%. Individual extraction efficiencies all fell within the range deemed acceptable by the US EPA, SW-846 method 8000B [US EPA, 1997] of 70-120%. The relative standard deviation was found to be no greater than 12 %, indicating satisfactory extraction and calibration repeatability.

A further repeatability test involved the extraction of six replicate cells containing coal tar sample 7. The cells were prepared and extracted in the same way as the PAH-spiked samples. The surrogate recoveries fell within the range of 76-97%, again within the bracket recommended by the US EPA. The relative standard deviation (RSD) of surrogate recovery was found to be below 10% for all deuterated surrogates, indicating satisfactory extraction repeatability. The small range in mean surrogate recovery, 91.7% for D8-naphthalene to 88.6% for D12-chrysene, indicates that molecular mass does not cause a significant decline in extraction efficiency. The repeatability experiments also show that the evaporation step performed after ASE (used to reduce and concentrate the extracts prior to analysis) does not give significant loss of analytes.

## 6.3.2   Preliminary Studies

Preliminary analysis of coal tar extracts was performed to evaluate the benefit of GCxGC over conventional techniques. GCxGC is a more expensive technique so if there was no great benefit it would not be economical to employ it for coal tar analysis. Figure 6.5 demonstrates the enhanced separation of coal tar constituents by GCxGC with respect to GC-MS. The single coal tar extract produced by ASE was analysed by both GC-MS and GCxGC, which separated around 300 and 3000 peaks respectively, thus proving that GCxGC is capable of separating an order of magnitude more compounds than conventional techniques. A single extract would not generally be analysed by GC-MS without tedious sample fractionation, but Figure 6.5 illustrates the great potential for fast screening of coal tars by ASE offline GCxGC, as it negates the requirement for such fractionation steps.
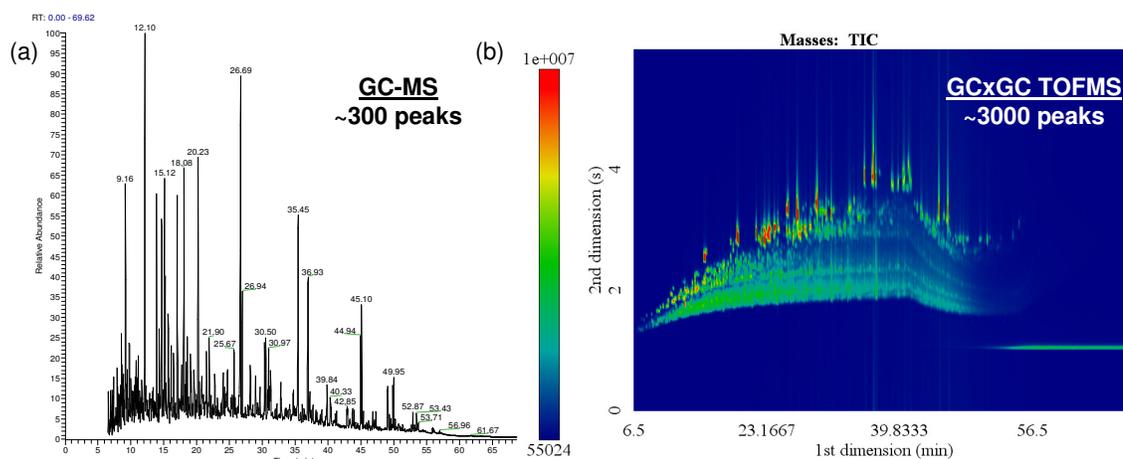
Figure 6.5: Comparison of the resolution power of (a) GC-MS and (b) normal phase GCxGC TOFMS through analysis of a coal tar extract (sample 12 in hexane).

### 6.3.3   Optimisation of Separation

GCxGC TOFMS was shown to have high potential for coal tar chemical fingerprinting, thus the next step was to optimise the instrumental parameters to allow as much chemical information to be obtained as possible. GCxGC separation is governed by a number of factors, as in conventional GC systems, however there are now two separations to consider. The temperature program, column choice (including column dimensions and stationary phase coating) and modulator settings all play vital roles in determining the analyte elution order and chromatographic resolution.

**Column Selection**

The first stage in optimisation involved the choice of column set. Initial analyses employed a normal phase column set consisting of a non-polar TR5MS (30 m x 0.25 mm i.d. x 0.25 $\mu$m film thickness) coupled to a mid-polarity TR50MS (1.5 m x 0.25 mm id x 0.25 $\mu$m film thickness) as supplied with the instrument. The contour plot provided in Figure 6.6 illustrates the major chemical classes identified in a typical coal tar by normal phase GCxGC. The broad peaks and poor separation in the second dimension suggests poor focussing of the first column effluent or excessive retention within the secondary column. The relatively wide internal diameter (0.25 mm) and thick stationary phase coating (0.25 $\mu$m) of the secondary column was thought to be responsible for the broad peaks.

146

Consequently, the column was replaced by a narrower, equivalent phase column, specifically an Rxi-17 (1.5 m x 0.1 mm id x 0.1 $\mu$m film thickness). As shown in Figure 6.7, this provided the desired effect, however, the maximum operating temperature of the secondary column was 260 °C which restricts the oven temperature program for elution of high boiling point compounds. It was found that exceeding the maximum operating temperature resulted in the column becoming brittle and easily snapped within the small confines of the secondary oven.
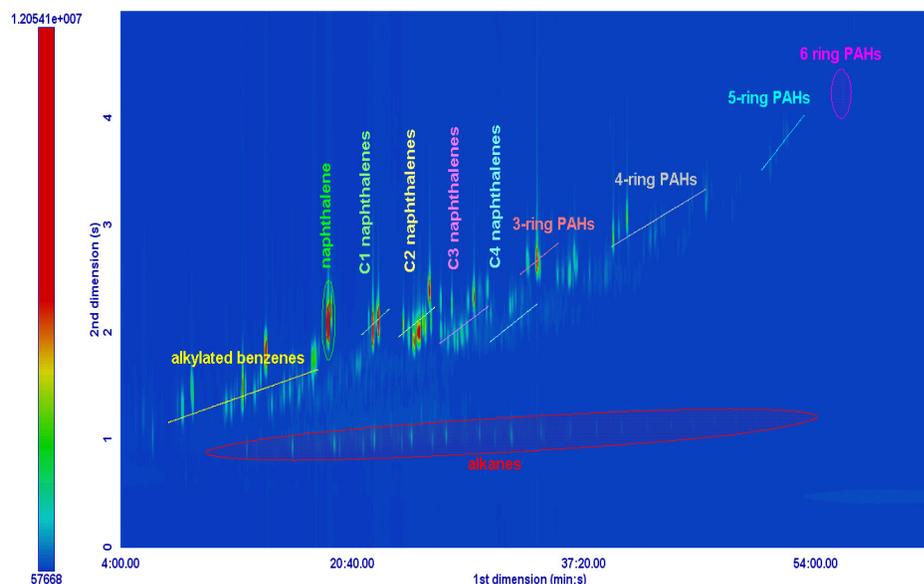


Figure 6.6: Initial classification of chemical classes in coal tar sample 1 using normal phase (column set 1 in Table 6.1) GCxGC.

A number of other mid-polar to polar columns were tested in the second dimension (see Table 6.1) in an attempt to increase orthogonality. However, similar effects were observed and columns became brittle very quickly. Additionally, the increased polarity of the DB-Wax and Rtx-200 columns caused fast elution of the relatively non-polar coal tar extracts in the second dimension. This resulted in contour plots with eluting compounds restricted to a one second band in the second dimension. Regardless of alterations to instrumental parameters, such as modulation temperature/time and secondary oven temperature, the second dimension separation could not exceed that of column set 2 which employed a narrow Rxi-17 column in the second dimension. However, the issue of exceeding the maximum operating temperature for this column still existed.
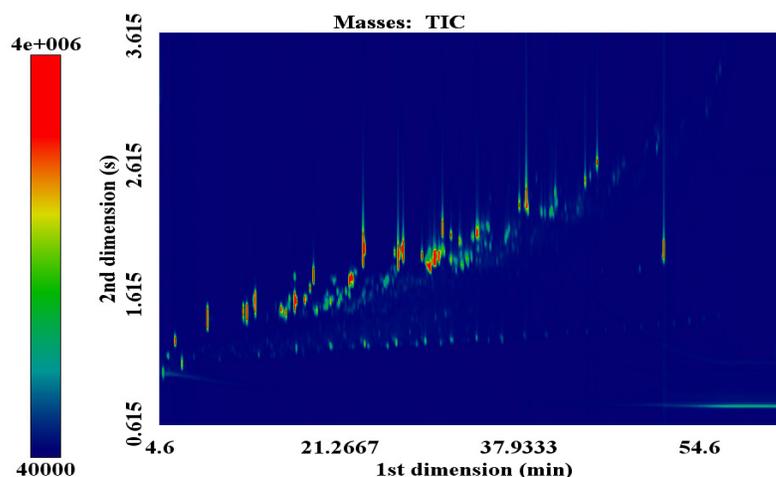
Figure 6.7: GCxGC contour plot illustrating the improved resolution achieved in normal phase mode by employing the narrower secondary column in column set 2 (as listed in Table 6.1).

For this reason, the column phase was reversed to allow elution of the high molecular weight compounds present in coal tars without damaging the secondary column, as non-polar columns are capable of withstanding higher secondary oven temperatures. A standard mixture of C2 alkyl naphthalenes was used to compare the separating power of three GC methods; GC-MS, normal phase GCxGC (column set 2) and reversed phase GCxGC (column set 5). Alkyl PAHs are often used in diagnostic ratios for source determination [Douglas et al., 2007]. However, due to insufficient separation by conventional GC techniques, the alkyl PAHs are generally combined based alkylation level, to provide ratios for the group as a whole. For example, a typical diagnostic ratio for alkyl naphthalenes would be N0/(N2+N3), where N0 represents the parent naphthalene while N2 and N3 are the C2 and C3 alkyl naphthalenes respectively [Douglas et al., 2007]. GCxGC is capable of separating the individual components making up each alkyl band to allow a more comprehensive and robust chemical fingerprint to be obtained. The chromatograms of the separation of a mixture of alkyl naphthalene isomers using GC-MS, normal phase GCxGC and reversed phase GCxGC are presented in Figure 6.8.

Table 6.1: Column sets evaluated for optimal separation of coal tars [P = polar and NP=non polar].

| Column Set # | Primary column | | | Secondary column | | |
|---|---|---|---|---|---|---|
| | column type and dimensions | stationary phase | polarity | column type and dimensions | stationary phase | polarity |
| 1 | TR5MS (30 m x 0.25 mm i.d. x 0.25μm) | 5% diphenyl/95% dimethyl polysiloxane | NP | TR50MS (1.5 m x 0.25 mm i.d. x 0.25μm) | 50% diphenyl/50% dimethyl polysiloxane | P |
| 2 | Rtx5SilMS (30 m x 0.25 mm i.d. x 0.25μm) | crossbond, selectivity close to 5% diphenyl/ 95% dimethyl polysiloxane | NP | Rxi-17 (1.5 m x 0.10 mm i.d. x 0.10μm) | 50% diphenyl/50% dimethyl polysiloxane | P |
| 3 | Rtx-5SilMS (30 m x 0.25 mm i.d. x 0.25μm) | crossbond, selectivity close to 5% diphenyl/ 95% dimethyl polysiloxane | NP | DB-Wax (1.5 m x 0.10 mm i.d. x 0.10μm) | 100% polyethyleneglycol | P |
| 4 | Rtx-5SilMS (30 m x 0.25 mm i.d. x 0.25μm) | crossbond, selectivity close to 5% diphenyl/ 95% dimethyl polysiloxane | NP | Rtx-200 (1.5 m x 0.10 mm i.d. x 0.10μm) | crossbond, 50% trifluoro-propylmethyl polysiloxane | P |
| 5 | TR50MS (30 m x 0.25 mm i.d. x 0.25μm) | 50% diphenyl/50% dimethyl polysiloxane | P | Rtx-5 (1.5 m x 0.18 mm i.d. x 0.20μm) | 5% diphenyl/95% dimethyl polysiloxane | NP |
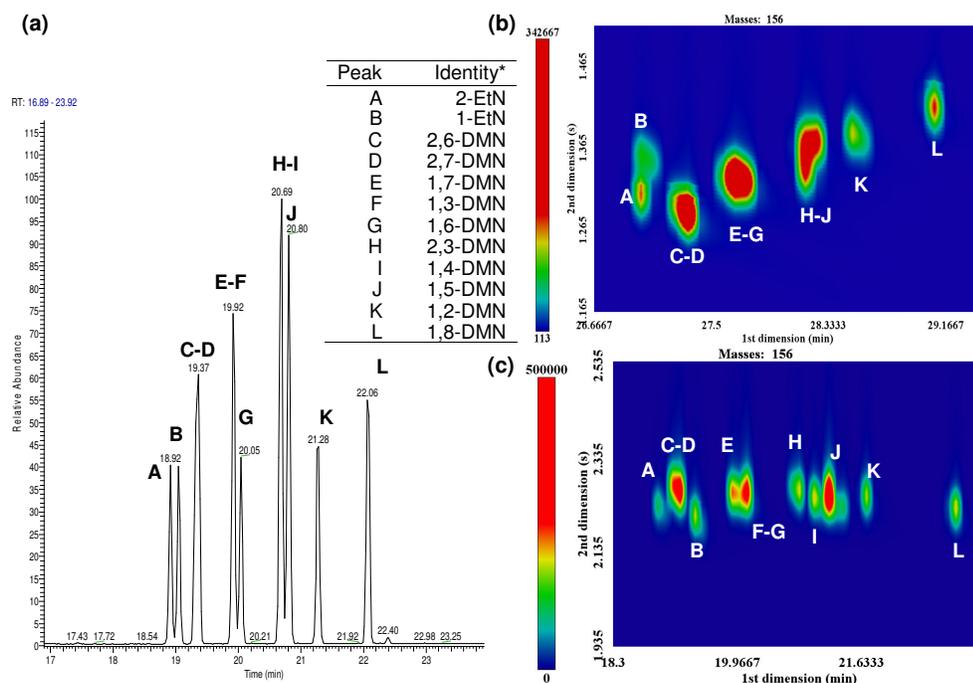
Figure 6.8: Comparison of the separation capabilities of (a) GC-MS, (b) normal phase GCxGC and (c) reversed phase GCxGC using a standard mixture of C2 alkyl naphthalene isomers. Peak identities are provided in the table (*EtN=ethyl naphthalene, DMN=dimethyl naphthalene) [reproduced from McGregor et al., 2011a].

Normal phase GCxGC and GC-MS achieved separation of 7 and 9 peaks respectively. Reversed phase GCxGC provided enhanced separation with only 2 pairs of the alkyl naphthalenes still co-eluting (2,6-and 2,7-dimethyl naphthalene and 1,3-and 1,6-dimethyl naphthalene). Surprisingly, GC-MS provided better resolution of the alkyl naphthalenes than normal phase GCxGC, which is generally employed in the analysis of complex mixtures. The heightened peak capacity of reversed phase GCxGC is further demonstrated in the C3 and C4 alkyl naphthalenes (Figure 6.9) as well as a host of other alkyl PAHs, such as the alkyl benzothiophenes (Figure 6.10). Reversed phase GCxGC separated 14 out of 34 possible C3 naphthalene isomers and 20 out of 112 possible C4 naphthalenes, while normal phase resolves only 9 and 14 isomers respectively within the same coal tar extract. It is worth mentioning that these comparisons are to the total number of possible isomeric structures, it is not known how many of these isomers are actually likely to be present in coal tar. Nevertheless, the reversal of column set provides information on eleven additional compounds within this class alone. The alkyl benzothiophenes chromatograms presented in Figure 6.10 show a similar gain in

resolution, with eight additional peaks resolved by the reversed phase column set. The additional chemical information is obtained at no extra cost using the same analytical instrument.
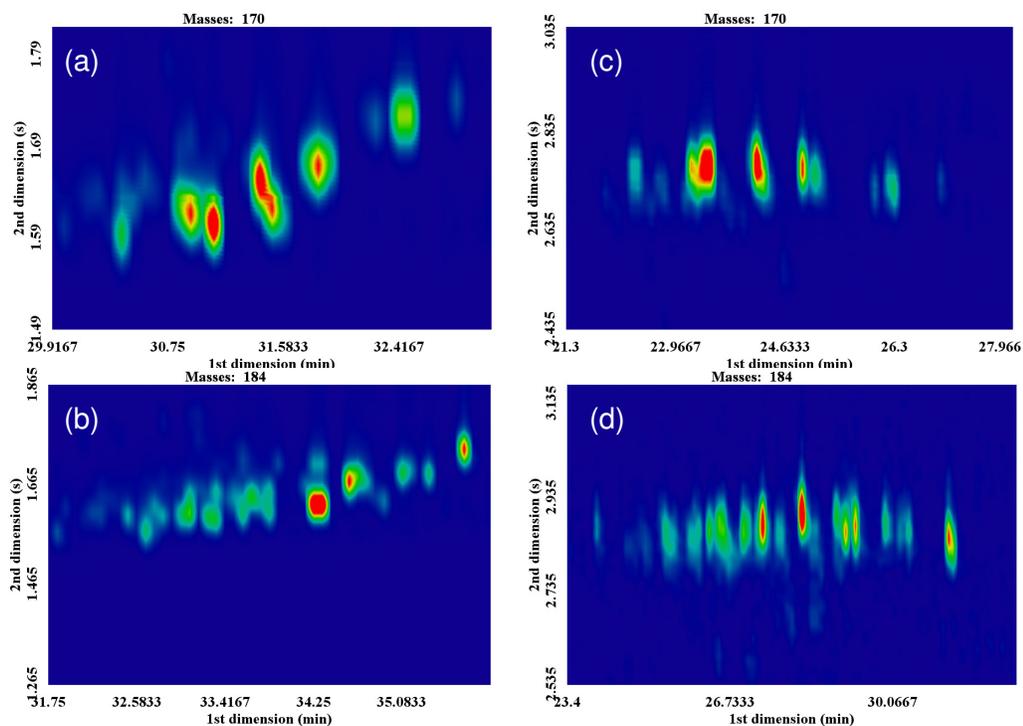


Figure 6.9: GCxGC contour plots of C3 and C4 alkyl naphthalenes (in coal tar sample 1) using normal phase GCxGC, (a) and (b) respectively, and reversed phase GCxGC, (c) and (d) respectively [reproduced from McGregor et al., 2011a].
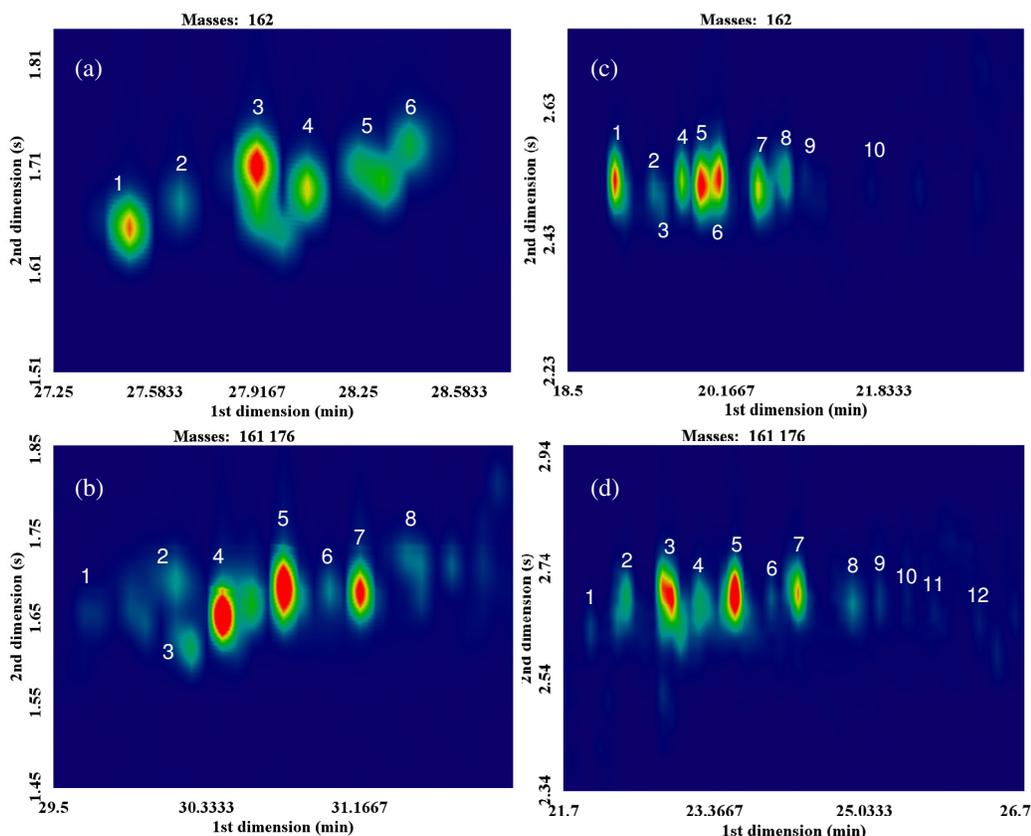
Figure 6.10: GCxGC contour plots of C2 and C2 alkyl benzothiophenes (in coal tar sample 1) using normal phase GCxGC, (a) and (b) respectively, and reversed phase GCxGC, (c) and (d) respectively. The numbering indicates the peaks identified as alkyl benzothiophene isomers by mass spectra [reproduced from McGregor et al., 2011a].

A full chromatogram of a coal tar extract produced by reversed phase GCxGC TOFMS is presented in Figure 6.11. The chromatogram illustrates the structured ordering caused by the two orthogonal separations. Compounds from the same chemical class elute together in bands, allowing fast, tentative identification of the major components present in the mixture (as shown by the annotations in Figure 6.11). Figure 6.11 also depicts the alternative elution order provided by reversed phase GCxGC. In normal phase GCxGC, the low affinity of aliphatics (including alkanes, alkenes and cycloalkanes) for the polar stationary phase in the second dimension results in the fast elution of these compounds and close clustering of the peaks in the lower region of the contour plot. On the other hand, the aliphatics elute after the aromatics in the second dimension in reversed phase, giving greater use of the chromatographic space in the upper regions of the contour plot.

152
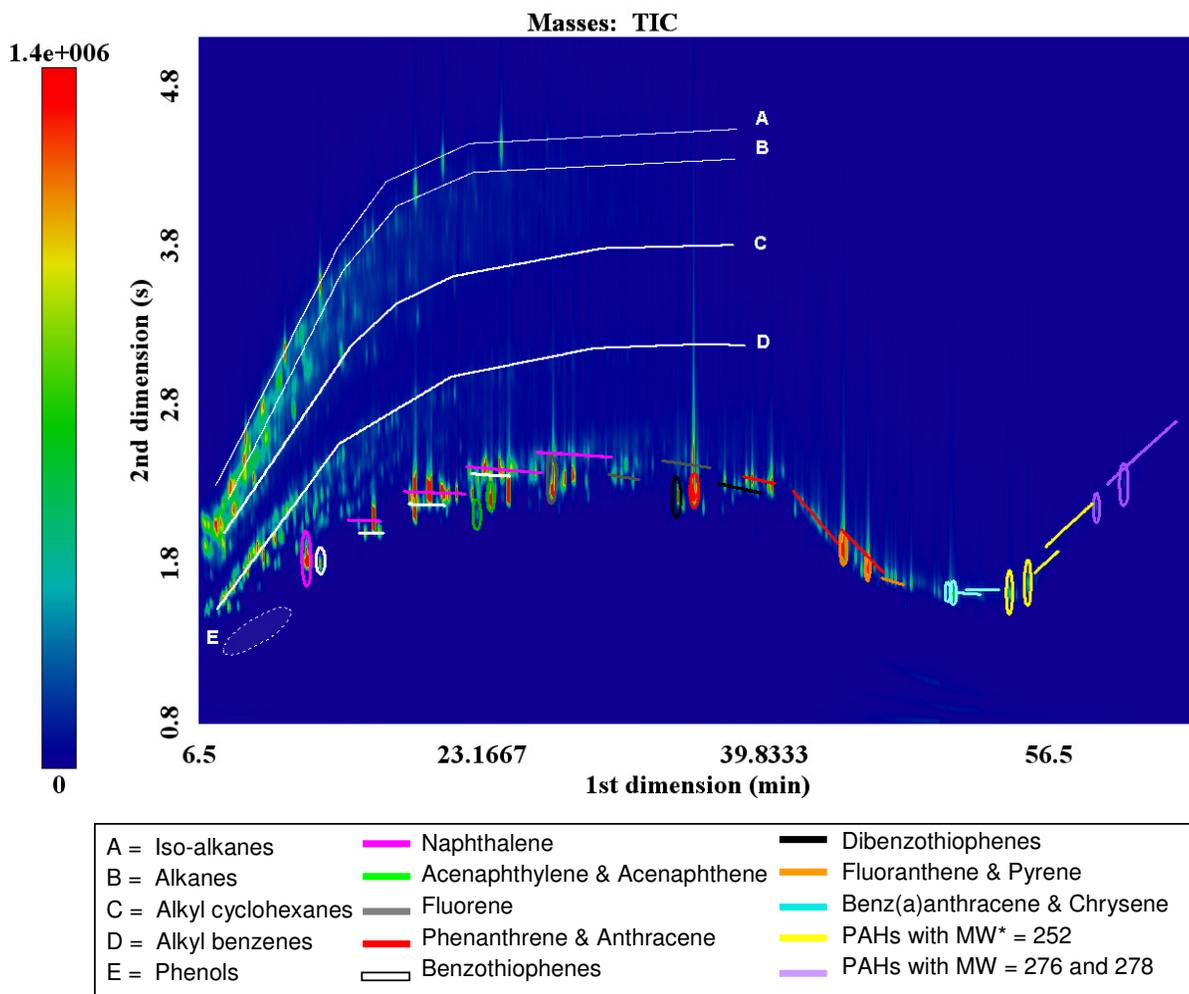
Figure 6.11: Contour plot (in total ion mode) of a coal tar extract analysed by reversed phase GCxGC TOFMS, where *MW = molecular weight. The solid lines of the same colour denote the alkylated derivatives of the corresponding circled compound [reproduced from McGregor et al., 2011a].

## Chromatographic Resolution and Space Occupation

The resolution of GCxGC chromatograms can be assigned a numerical value in a similar way as in conventional one-dimensional chromatography. As described in Chapter 3, the resolution of two peaks in a one dimensional chromatogram is measured by:

$$R = \frac{2d}{W_A + W_B} \tag{6.3.1}$$

where $d$ is the retention time difference between two peaks and $W_A$ and $W_B$ are the peak widths for the compounds A and B respectively.

On the other hand, for GCxGC resolution, the peak width and retention time in both dimensions must be considered, as illustrated in Figure 6.12. The resolution of GCxGC plots can be calculated by:

$$R_{2D} = \sqrt{\frac{2\left(\Delta_{tr}^{1}\right)^2}{\left(^1W_A +^1 W_B\right)^2} + \frac{2\left(\Delta_{tr}^{2}\right)^2}{\left(^2W_A +^2 W_B\right)^2}} \qquad (6.3.2)$$

where $^1W_A$ and $^2W_A$ are the peak widths of compound $A$ in the first and second dimensions respectively, while $^1W_B$ and $^2W_B$ are the peak widths of compound $B$ in the first and second dimensions respectively. $\Delta_{tr}^{1}$ and $\Delta_{tr}^{2}$ are the differences in retention times between the apexes of compounds $A$ and $B$, in the first and second dimensions respectively [Omais et al., 2011].
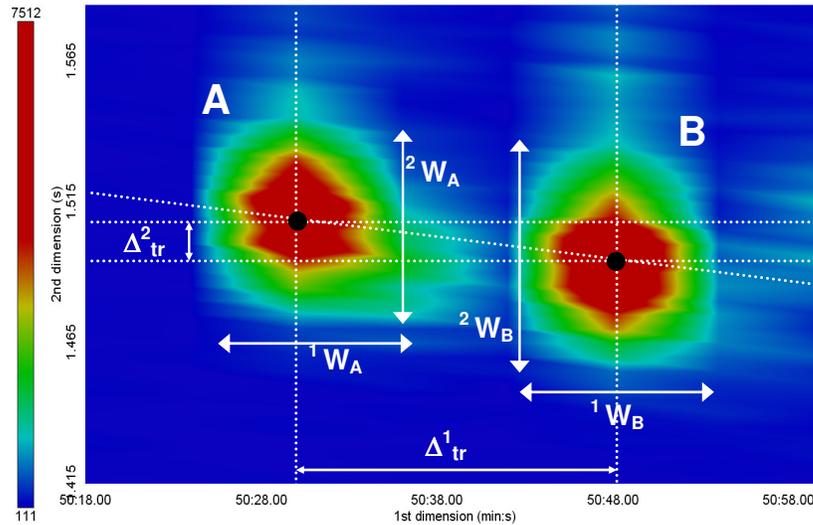


Figure 6.12: Schematic illustrating the variables used within the 2D resolution ($R_{2D}$) calculation shown in Equation 6.3.2 (adapted from Omais et al., 2011).

Equation 6.3.2 was used to compare the resolving power of the optimal normal and reversed phase column sets for a range of isomer combinations; the results are presented in Figure 6.13. The results show that reversed phase (column set 5) produced better resolution than normal phase (column set 2) between all the investigated isomers.
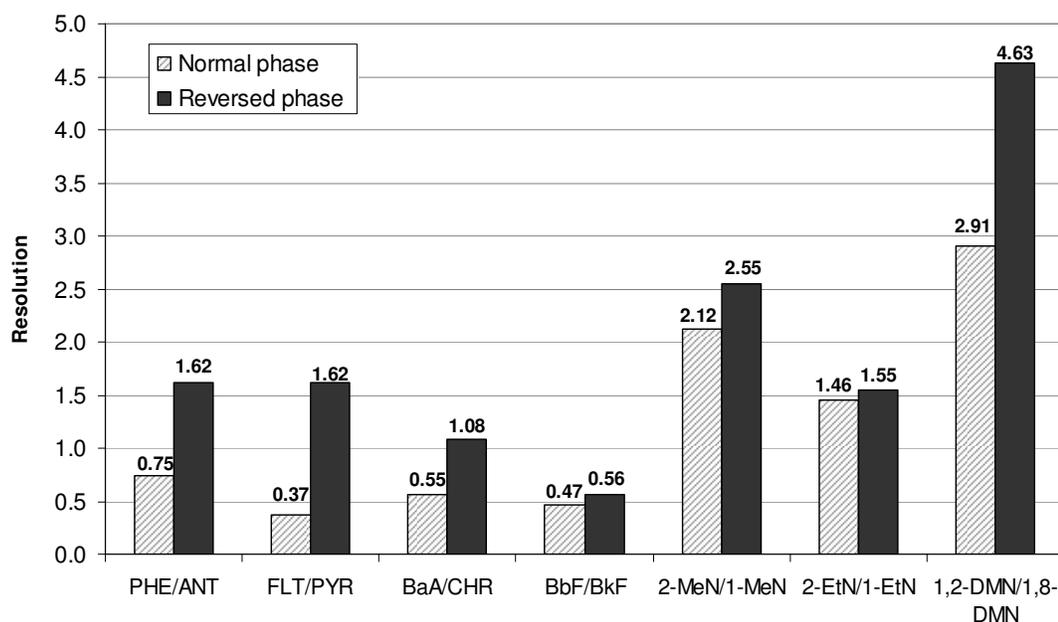
Figure 6.13: Comparison of resolutions for a number of PAH isomer pairs using both normal and reversed phase configurations (column sets 2 and 5 respectively).

The chromatographic space occupation within a GCxGC contour plot can be estimated by the following:

$$SO_{2D} = \sqrt{\frac{(^1tr_n - \, ^1tr_1)^2}{2t_{analysis}^2} + \frac{(^2tr_n - \, ^2tr_1)^2}{2t_{mod}^2}} \text{ x } 100 \qquad (6.3.3)$$

The space occupation ($SO_{2D}$) can be presented as a percentage using the difference in retention times of the first eluting compound ($tr_1$) and last eluting compound ($tr_n$) in both dimensions, the total analysis time ($t_{analysis}$) and the modulation time ($t_{mod}$) [Omais et al., 2011]. Equation 6.3.3 was used to compare the space occupation when using both normal and reversed phase column sets. It was found that normal phase utilised 76.2% of the available chromatographic space, while the reversed phase column set provided an improvement in orthogonality by utilising 80.5% of the total space.

155

### 6.3.4 Repeatability

The reversed phase GCxGC TOFMS method was validated through analysis of six replicate injections of a coal tar extract (sample 8). Quantification of the 16 EPA PAHs within the extract was compared using three GC methods; GC-MS, normal phase GCxGC and reversed phase GCxGC. The results are given in Table 6.2 and presented graphically in Figure 6.14. The poor resolution of GC-MS is evident in the large relative standard deviations (RSD) obtained for the higher molecular weight compounds. This effect is emphasised in closely eluting compounds, such as the benzofluoranthene isomers, which exhibit the highest RSD values of all the analysed PAHs. The coelution of such peaks causes poor accuracy and repeatability of quantification. On the other hand, both GCxGC methods exhibit improved repeatability for PAH quantification, with all RSD values below 10% for normal phase and below 5% for reversed phase.
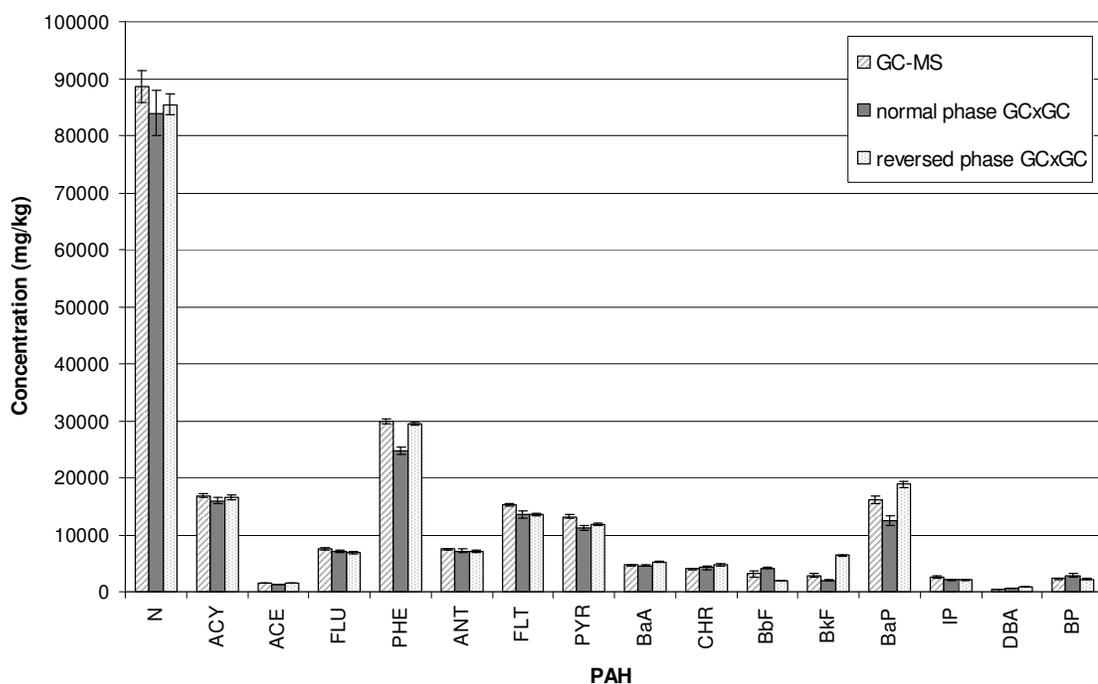


Figure 6.14: Comparison of gas chromatographic method repeatability for the quantification of the EPA priority PAHs found in coal tar samples. Values as listed in Table 6.2.

Table 6.2: Comparison of PAH quantification values (for coal tar sample 8) obtained using GC-MS, normal phase GCxGC TOFMS and reversed phase GCxGC TOFMS.

| Compound Name | GC-MS method | | normal phase GCxGC | | reversed phase GCxGC | |
|---|---|---|---|---|---|---|
| | Mean Concentration (mg/kg) | RSD[b] (%) | Mean Concentration (mg/kg) | RSD[b] (%) | Mean Concentration (mg/kg) | RSD[b] (%) |
| naphthalene | 88535 | 3.2 | 83960 | 4.8 | 85432 | 2.2 |
| acenaphthylene | 16865 | 2.4 | 16017 | 3.4 | 16567 | 2.7 |
| acenaphthene | 1510 | 2.9 | 1301 | 1.5 | 1485 | 3.9 |
| fluorene | 7509 | 3.1 | 7172 | 3.3 | 6970 | 3.3 |
| D10-phenanthrene (IS) | N/A | 0.8 | N/A | 0.4 | N/A | 0.3 |
| phenanthrene | 29902 | 1.5 | 24689 | 2.8 | 29445 | 0.8 |
| anthracene | 7445 | 2.3 | 7193 | 3.7 | 7194 | 3 |
| fluoranthene | 15271 | 1.7 | 13548 | 4.7 | 13630 | 1.4 |
| pyrene | 13224 | 1.7 | 11192 | 3.6 | 11821 | 1 |
| benz[a]anthracene | 4651 | 3.3 | 4608 | 3.1 | 5184 | 2.3 |
| chrysene | 4022 | 2.2 | 4207 | 6.8 | 4759 | 3 |
| benzo[b]fluoranthene | 3175 | 17.6 | 4187 | 3.7 | 1957 | 4 |
| benzo[k]fluoranthene | 2874 | 9.3 | 1972 | 4.1 | 6408 | 2.2 |
| benzo[a]pyrene | 16102 | 4 | 12546 | 6.7 | 18823 | 2.6 |
| indeno[1,2,3-cd]pyrene | 2609 | 5.4 | 2065 | 4.8 | 2133 | 4.5 |
| dibenz[a,h]anthracene | 423 | 8.1 | 570 | 2.6 | 913 | 4.2 |
| benzo[g,h,i]perylene | 2267 | 8.2 | 2899 | 9.0 | 2183 | 4.8 |
| Surrogate | Mean Recovery (%) | RSD (%) | Mean Recovery (%) | RSD (%) | Mean Recovery (%) | RSD (%) |
| D8-naphthalene | 88.5 | 3.3 | 70.2 | 4.1 | 71.1 | 2.2 |
| D10- fluorene | 92.2 | 3 | 98.9 | 3.3 | 81.0 | 2.8 |
| D10-fluoranthene | 89.1 | 2.9 | 82.9 | 3.5 | 78.9 | 4.9 |
| D12-chrysene | 81.2 | 3.8 | 82.2 | 4.5 | 80.0 | 2.4 |

[a] IS = internal standard (150 $\mu$g/mL spike)

[b] RSD = relative standard deviation based on 6 replicate injections of coal tar sample 8.

### 6.3.5    Chromatographic Pattern Recognition

The characteristic chromatographic patterns of petroleum products, such as diesel, are well-established in literature. In conventional GC analyses, the high aliphatic content in diesel provides a signature bell-shaped hump of n-alkanes which is easily recognisable. However, it was not known how these mixtures would respond to reversed phase GCxGC analyses. Therefore, fresh samples of petrol, diesel and a mix of petrol and diesel (50:50 w/w) were extracted in the same way as the coal tars and analysed by GCxGC TOFMS. Petrol and diesel samples were obtained fresh from a petrol station (in Stevenston, North Ayrshire) and kept refrigerated at 4 °C in sealed containers prior to analysis. The analysis of these samples provided a small set of reference petrogenics for comparison with the pyrogenic coal tar library.

The initial stage of data analysis involved the simple comparison of coal tar GCxGC chromatograms. The enhanced separation capacity of GCxGC provides large volumes of chemical data. In each case, over one thousand peaks were identified by the instrument software (ChromaTOF®, version 4.22, Leco) with certain samples containing over 3000 different components. Therefore, the visual comparison of chromatograms is often a time-consuming, cumbersome process where only the major differences between samples are easily noticeable.

However, in this work certain samples which were produced by different manufacturing processes gave distinctly different GCxGC chromatograms. For example, CWG tars generally presented a high content of low molecular weight compounds, such as the straight, branched and cyclic alkanes, while the fresh coke oven tars were dominated by parent PAHs. The comparison of four coal tar samples is shown in Figure 6.15; the chromatograms show that while certain coal tars from different manufacturing processes can be distinguished, identifying chromatographic differences between tars produced by the same manufacturing process is very difficult.

Due to the high concentrations of PAHs and alkyl PAHs in most samples, smaller peaks can be masked in the chromatograms, causing comparison on a visual basis alone to be challenging, especially when the identity of the relevant peaks is unknown. The subjectivity and tedious nature of chromatographic pattern recognition highlights the need for an alternative method of sample comparison. For this reason, a number of chemometric methods, which apply mathematic principals for the comparison of chemical data, will be evaluated in the following Chapter.
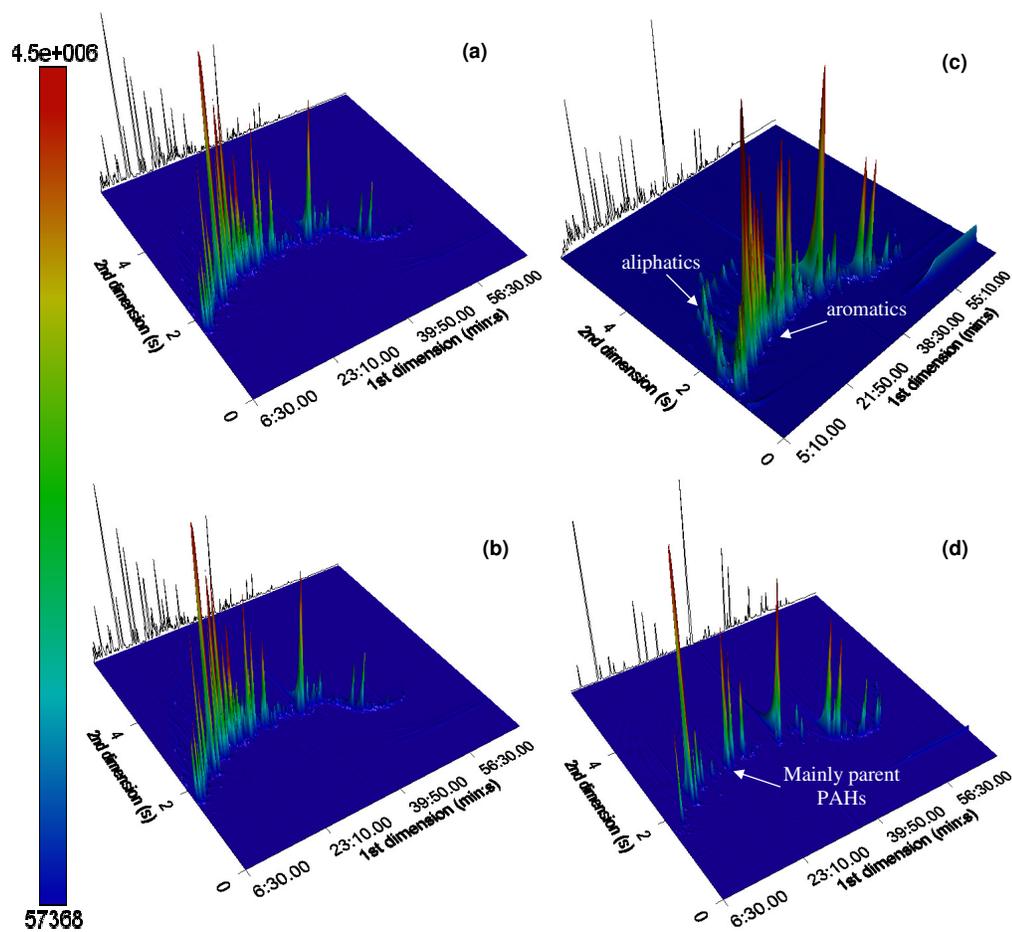
Figure 6.15: Comparison of GCxGC surface plots in total ion mode for two CWG tars (a) and (b) [samples 13 and 14 respectively], a creosote oil (c) [sample 11] and a fresh coke oven coal tar (d) [sample 18]. The back plane represents the reconstituted one-dimensional chromatogram. Reproduced from McGregor et al., 2011b.

## 6.4   Conclusion

This Chapter has illustrated the great potential of ASE offline reversed phase GCxGC for ultra resolution chemical fingerprinting of coal tar. GCxGC has the capacity to separate an order of magnitude more compounds than conventional GC. Simple reversal of GCxGC column polarity was shown to further increase the resolution power, with no added cost to the system. Conventional tiered approaches for chemical fingerprinting of complex mixtures involve time-consuming sample preparation and clean-up steps and analyses on multiple analytical instruments. The enhanced separation capacity allows a range of chemical classes to be monitored in a single-step analytical procedure, negating the requirement for tedious sample preparation and clean-up protocols. The ASE offline reversed phase GCxGC method couples a highly efficient, automated extraction procedure with accurate and precise chromatographic separation of an entire sample in a single analytical run; deemed ultra resolution chemical fingerprinting.

However, the large volumes of chemical data produced by GCxGC causes difficulties in the comparison of complex samples. Major compositional differences can be identified by visual comparison of chromatograms, yet it is easy to overlook compounds which may hold vital information for source provenancing. The following Chapter investigates the comparison of GCxGC datasets using a variety of statistical methods, with the goal of developing a classification model capable of source-specific allocation of coal tars.

# Chapter 7

# Source Apportionment of Coal Tar

## 7.1 Introduction

The introduction of environmental legislation such as the European Liabilities Directive (2004/35/EC) which promotes the polluters pay policy has instigated the requirement for accurate source allocation of environmental contaminants.

Following the decline of the manufactured gas industry in the mid-twentieth century, many gas plants were demolished, redeveloped and often split into multiple land holdings. Heavy coal tar contamination is a common occurrence at former manufactured gas plants thus it is important to be able to allocate coal tar to its formation by a particular process. Due to the high abundance of contaminants, such as PAHs, formed as by-products of the gas manufacturing industry, it is easy to attribute any environmental contamination in the vicinity of a gas plant to the past manufacturing processes which would have occurred on site without a thorough forensic investigation.

The previous Chapter presented a new accurate and precise method of chemical fingerprinting by ASE offline reversed phase GCxGC. However, the GCxGC data is meaningless without detailed interpretation, as chromatographic pattern recognition tends to involve a degree of subjectivity. The large volumes of data obtained by GCxGC calls for robust statistical comparison techniques. Statistical techniques are generally classified as either univariate or multivariate. As the name suggests, univariate statistics involve the evaluation of one variable at a time, for example, cross plots
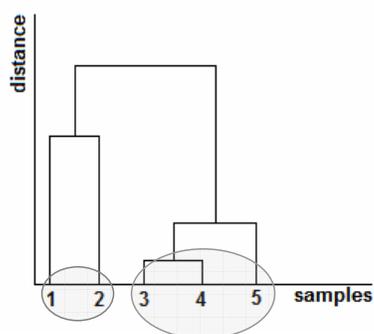
Figure 7.1: Example of an HCA dendrogram.

illustrating diagnostic ratios. On the other hand, multivariate statistics are capable of evaluating numerous variables simultaneously in an effort to condense large datasets into a more manageable format.

### 7.1.1   Chemometric Techniques

Chemometrics is the application of statistical or mathematical techniques to chemical data [Christensen and Tomasi, 2007]. There are a great number of chemometric techniques which may be employed for the interpretation of GC data, but only the techniques relevant to this work will be discussed; namely hierarchical cluster analysis (HCA) and principal component analysis (PCA). These techniques are multivariate statistical methods which are capable of exploring correlations between large datasets, with the relevant information represented in a simple graphical form. It must be emphasised that multivariate statistics do not provide a single, correct answer; they are exploratory techniques and it is the analyst's responsibility to interpret the results.

### Hierarchical Cluster Analysis

Hierarchical cluster analysis (HCA), as the name suggests, does not involve single-step classification of data, the aim is to produce a hierarchical set of linkages between similar objects. A number of partitioning stages are used to produce the final groupings. HCA classifications can be viewed as a two-dimensional graph, known as a dendrogram, which illustrates the linkages made at each stage of the analysis.

For example, the dendrogram in Figure 7.1 shows that samples 3 and 4 are most similar and two distinct clusters are evident (highlighted by the shaded circles) within the dataset. HCA can identify the main clusters within a dataset using two principal methods; agglomerative clustering or divisive clustering. The more popular agglomerative method begins with all observations treated as separate entities. The observations are then gradually clustered based on their degree of similarity until all clusters are joined as a single group. On the other hand, the divisive method initially treats all observations as a single cluster which then branches out into smaller groups until each branch represents a single observation. However, in the divisive method, once assignment to a cluster has been performed the observation cannot be assigned to an additional cluster in a following tier. Only agglomerative clustering was performed in this study.

**Distance measurement**     The degree of similarity between observations (or variables) is measured by their proximity to one another. There are numerous methods of distance measurement, including Euclidean, Manhattan and Manalonobis, however for the purpose of this study only Euclidean distance will be discussed. The measurement of Euclidean distance can be represented by the following:

$$d_{AB} = \left[ \sum_j \left( x_{1,j} - x_{2,j} \right)^2 \right]^{\frac{1}{2}} \tag{7.1.1}$$

where $d_{AB}$ represents the distance between two observations, A and B, and $x_{i,j}$ is the value of the $j^{th}$ variable measured on the $i^{th}$ object.

**Linkage Mechanism**     An additional parameter involved in HCA, known as the linkage mechanism, is required to describe how the distance between clusters containing multiple members is measured. The Euclidean distance measures the distance between two data points, therefore, a defined procedure on how to link each group must be established. A number of different linking mechanisms exist; with single, average and complete linkage being the most widely used.

The various linkage mechanisms are illustrated in Figure 7.2. Single linkage uses the minimum distance between two clusters, while complete linkage uses the maximum distance. Finally, average linkage uses the average distance between all sample pairs in the two clusters. It is standard practice
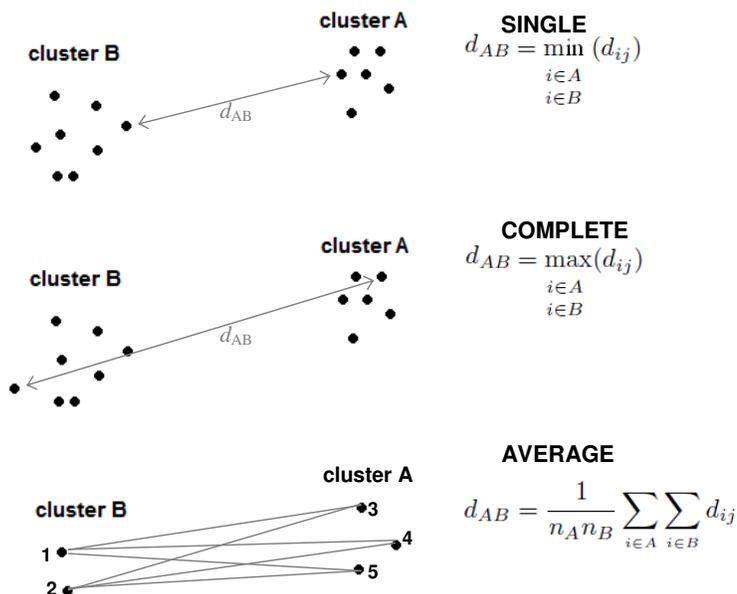
Figure 7.2: Linkage mechanisms employed in hierarchical cluster analysis (adapted from Everitt and Hothorn, 2011).

to compare a number of linkage mechanisms for a given dataset, as they may provide significantly different clustering options, allowing the optimal results to be obtained.

**Two-dimensional Hierarchical Cluster Analysis**   Two-dimensional HCA simultaneously evaluates the rows and columns of a data matrix to produce clustering of both samples and variables. This generates a heatmap or clustergram consisting of two separate dendrograms connected by a tiled mosaic. The tiling connects the variable and sample clusters and is colour-coded according to the intensity of the variable within that particular sample [Wilkinson and Friendly, 2009]. Heatmaps are well-established within bioinformatics for visualising genome expression patterns, only recently has it been applied for clustering contaminated samples based on their chemical composition [Fernandez Varela et al., 2010]. As with conventional HCA, the linkage and distance mechanisms can be varied to achieve dendrograms which best describe the dataset.

164

**Principal Component Analysis**

Principal component analysis (PCA) is an additional method of reducing large datasets to better visualise trends and patterns within the data. A dataset (X) containing $p$ variables in columns and $n$ samples or objects in rows, can be represented in $p$-dimensional space where each axis represents a different variable (Figure 7.3). Lines representing the directions of greatest dispersion of $n$ points can then be added. The rotation of the original co-ordinate system onto the new orthogonal axes, known as principal axes, can illustrate the maximum variation in the total dataset; this forms the basis of principal components analysis (PCA). The original data points become projected onto the new axes, for example, point $(x_{1m}, x_{2m})$ shown in Figure 7.3, projects to point $Y_{1m}$ on the the $Y_1$ axis and $Y_{2m}$ on the $Y_2$ axis. The system is such that the projected point $Y_{1m}$ has a value of $m = 1...n$ which is greater than if the point was projected onto any other axis passing through the mean of the dataset $(\overline{x}_1, \overline{x}_2)$ [Everitt and Dunn, 2001].
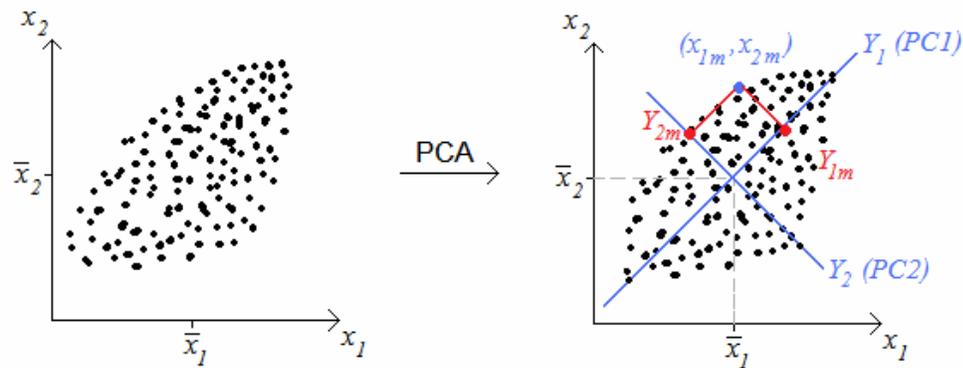


Figure 7.3: Projection of principal axes onto a dataset illustrating the directions of greatest variation within the data.

When performing principal component analysis, the principal axes are referred to as principal components, or PCs. The goal of PCA is to produce a set of PCs which significantly reduce the original $p$ variables and account for nearly all the total variance of a dataset. Principal components are ordered in terms of decreasing variance; PC1 represents the most variation in the dataset, PC2 represents the next highest variance and so on [Everitt and Hothorn, 2011]. The subsequent PCs are orthogonal to the previous and are formed such that they explain the maximum amount of remaining variance in the dataset . The process is continued until the last $p^{th}$ principal component, but generally 2-3 PCs will sufficiently explain the total variance within a dataset [Everitt and

Hothorn, 2011]. The algebraic formulae used to derive the PCs are discussed in the following section.

**PCA Algebra**    Principal components are linear combinations of the original variables and are weighted according to their contribution to explaining the variance in a certain dimension. Equation 7.1.2 represents the linear combination used to create the first PC.

$$Y_1 = a_{11}x_1 + a_{12}x_2 + ...a_{1p}x_p = \mathbf{a}_1'\mathbf{x} \qquad (7.1.2)$$

where $a_i$ represents coefficients used to weight the variables and $i = 1..p$ [Everitt and Hothorn, 2011]. These combinations allow for the variance of $Y_1$ to be increased infinitely by increasing the value of the coefficients $a_{11}$, $a_{12}...a_{1p}$, which can also be written as the vector $\mathbf{a}_1$[Everitt and Dunn, 2001]. Therefore, restrictions must be applied. The condition for the first PC is that the sum of squares of the original data must equal 1, written as:

$$\sum_{i=1}^{p} a_i^2 = 1 \qquad (7.1.3)$$

The linear combination of data points to create the second principal component then explains maximum dispersion of data points within $(p-1)$ dimensions. The second PC must satisfy the conditions $\mathbf{a}_2'\mathbf{a}_2 = 1$ and $\mathbf{a}_2'\mathbf{a}_1 = 0$ to ensure that the two PCs are linearly independent [Everitt and Dunn, 2001]. Thus, for the $i^{th}$ PC the linear combination is shown by:

$$Y_i = a_{1i}x_1 + a_{2i}x_2 + a_{3i}x_3 + ...a_{ni}x_p \qquad (7.1.4)$$

The direction of the weights $(a_i)$ used in the linear combinations are given by the eigenvector corresponding to the highest eigenvalue in the covariance matrix ($\mathbf{S}$). Eigenvalues ($\lambda$) represent the variance of each principal component, allowing the PCs to be arranged in order of decreasing variance. The variance of a linear composite $\sum_{i=1}^{p} a_i x_i$ is given by:

$$Var(Y_1) = Var(\mathbf{a}_1'\mathbf{x}) = \mathbf{a}_1'\mathbf{S}\mathbf{a}_1 \qquad (7.1.5)$$

where $\mathbf{a}_1$ is the eigenvector of $\mathbf{S}$ corresponding to the largest eigenvalue [Everitt and Dunn, 2001]. The eigenvectors ($\gamma$) and eigenvalues ($\lambda$) of the covariance matrix, $\mathbf{S}$, are formed such that $\mathbf{S}\gamma=\lambda\gamma$. Thus, the relevant variance explained by each principal component is generally expressed as a percentage by:

$$\% \text{ explained variance} = \frac{\lambda_i}{\sum \lambda_i} \times 100 \tag{7.1.6}$$

A plot of eigenvalues, known as a scree plot, can be compiled to show the ability of each PC to describe the variance in a dataset. The final output of PCA is represented by a score plot, where the scores of each sample/object are plotted for the PCs which explain the most variance. The closer together the samples cluster, the more alike they are in character and vice versa. The weights (or loadings) applied to each variable within each PC can then be viewed to show the variables which are most important in deriving a particular PC.

## 7.1.2   Limitations of Chemometrics

Chemometric methods, such as PCA and HCA, can be powerful exploratory tools, however, without specific knowledge of the samples under investigation, it is possible to misinterpret the resulting classifications. For example, within environmental forensics, weathering is a particular problem and can cause degradation of samples which may significantly affect statistical groupings.

Moreover, noisy, irrelevant variables can obscure other variables which are essential for accurate classification [Fernandez Varela et al., 2010]. The careful selection of variables to be included in a statistical model is vital as the most important variables are not always obvious.

## 7.1.3   Chemometrics in Environmental Forensics

As mentioned in Chapter 2, the last stage of the traditional tiered approach to chemical finger-printing is the statistical analysis of data. Numerous chemometric techniques have been used in the field of environmental forensics for pattern recognition and classification purposes. PCA and HCA are well-established techniques for data analysis, however, it is only recently that they have

167

been adopted in environmental forensics [Christensen et al., 2004]. The complex nature of most environmental samples requires a reduction in dimensionality of data without any loss in relevant information; this can be achieved through multivariate statistical methods. Defensibility of a fingerprinting method requires the relevant information to be extracted to allow samples to be compared in an objective manner [Christensen and Tomasi, 2007].

PCA has been used within the field of geochemistry since the 1980s for oil-oil and oil-source rock correlations. The technique was first applied to environmental forensics in a publication by Aboul Kassim and Simoneit [1995] detailing the use of a multivariate statistical approach to differentiate between harbour sediment contaminant sources using aliphatic hydrocarbon data.

Christensen and Tomasi [2007] state that the use of a multivariate approach which avoids peak identification and quantification could prove most cost effective. For example, PCA can be applied to chromatographic peak data without time-consuming peak identification steps. Christensen et al. [2005] developed a method based on this theory, by combining GC-MS analyses, data preprocessing and PCA to provide a fast and effective form of source discrimination. Preprocessing of the data was shown to remove any variations unrelated to sample composition [Christensen et al., 2005]. For example, instrumental sensitivity changes or concentration effects can be minimised.

Fernandez Varela et al. [2010] employed a suite of chemometric techniques to investigate a set of diagnostic ratios for crude oil fingerprinting. Fernandez Varela et al. [2010] concluded that simple univariate/bivariate methods were not adequate for source differentiation and only provided vague trends between oil classes. The study showed that a suite of twenty-eight diagnostic ratios could be reduced to four essential variables without loss of classification power in the PCA scoreplot; this illustrates the importance of variable selection in statistical analyses. The potential of HCA heatmaps for environmental forensics was also demonstrated by Fernandez Varela et al. [2010], as applied to the clustering of oil samples based on their geographic origin. The technique was shown to provide strong classification of oils, with the added benefit of easy visualisation of the variable groupings and how they influence sample classification (Figure 7.4).
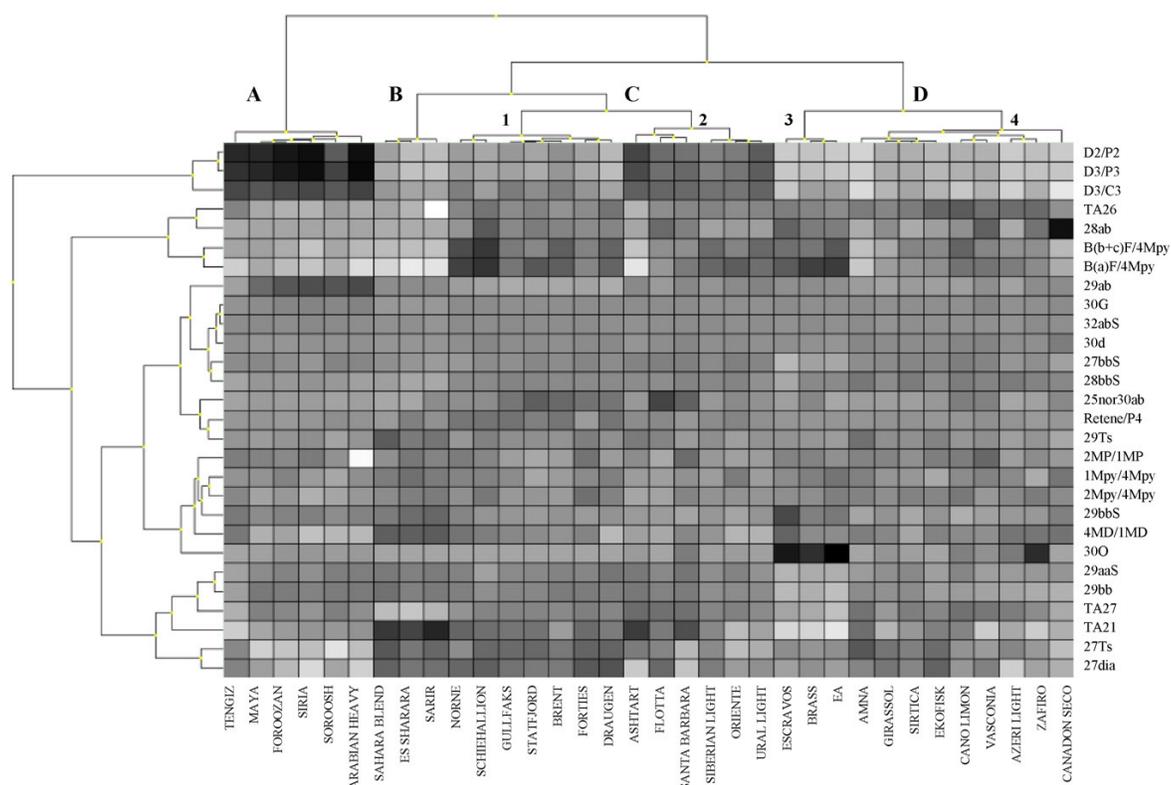
Figure 7.4: HCA heatmap of geographic oil classifications using a suite of 28 common diagnostic ratios, where clusters A-D represent groupings according to geographic location. A=Middle East, B=North Africa, C1=North Sea, C2=different origins, D3=Nigerian and D4=different origins (copied from Fernandez Varela et al., 2010).

Chemometric techniques have previously been applied for the interpretation of GCxGC data [Van Mispelaar et al., 2005]. The large quantities of information obtained through GCxGC analyses make multivariate methods an essential part of data processing. Van Mispelaar et al. [2005] employed PCA to detect chemical differences between 14 oil samples from three different reservoirs on the same oil field. In total, 65 variables were selected for PCA and the resulting score plot is given in Figure 7.5. Three dense clusters were produced corresponding to the three reservoirs sampled in the study. The outliers may be explained by the small data subset used for PCA (65 out of 3904 peaks) which may not be representative of the entire sample. The scoreplot clearly illustrates the classification power of GCxGC plus PCA, especially given that all oils in the study were sampled from the same oil field, thus would be expected to be very similar in chemical composition.
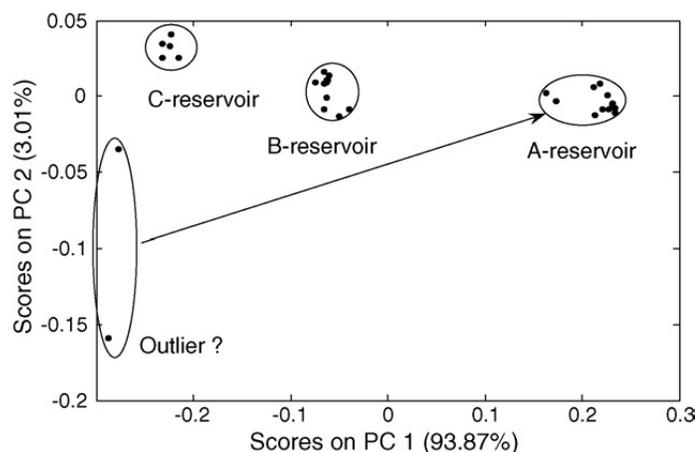
Figure 7.5: A PCA score plot produced by GCxGC datasets from the analysis of oil samples from three reservoirs within a single oil field (copied from Van Mispelaar et al., 2005).

A further example of PCA can be demonstrated through analysis of the carbon isotope dataset presented in Chapter 5 of this thesis. The resulting PCA score plot (Figure 7.6) shows a large number of the coal tar samples cluster together indicating similar PAH isotopic signatures. However, sample 11 is positioned slightly outside of this main cluster. This sample represents a distilled fraction of coal tar (namely creosote oil), thus the PAHs have been exposed to additional chemical reactions during distillation. Isotopic fractionation processes during these reactions are likely to have altered the isotopic signature of sample 11. Additionally, samples 2, 3, 10 and 15 are positioned well outside of the main coal tar cluster. These samples exhibited more negative $\delta^{13}$C values than the majority of coal tars, thus it is possible that they represent a mixture of petrogenic and pyrogenic sources. However, a larger database of PAH sources would be required to confirm this, which is outside the scope of this thesis. Figure 7.6 demonstrates the ability of PCA to simplify data and highlight the main trends within a dataset.
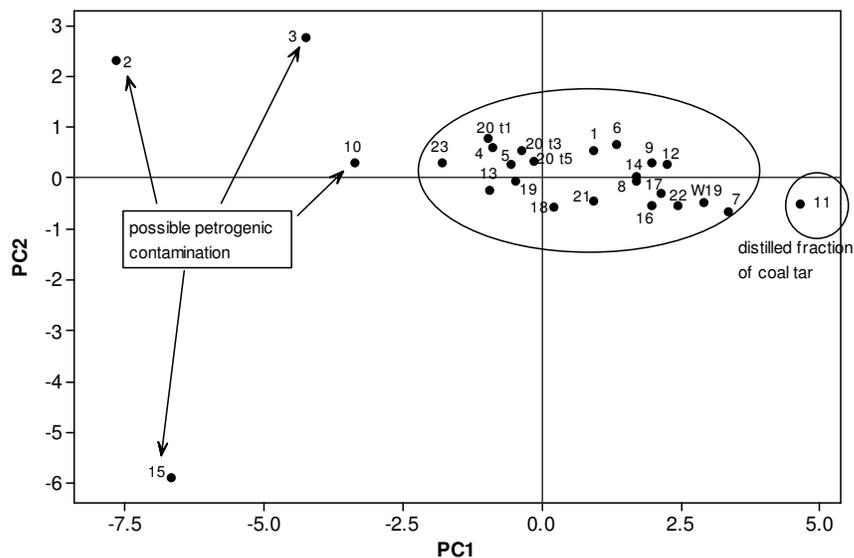
Figure 7.6: A PCA score plot using data produced by compound specific carbon isotope analysis of coal tar PAHs as detailed in Chapter 5.

To the author's knowledge, the outputs of this project represent the only publications investigating multivariate statistical analysis of coal tars using GCxGC datasets [McGregor et al., 2011a,b]). Previous publications have focussed solely on univariate methods of source apportionment, generally using GC-MS analyses. For example, Mauro [2000] presented the use of PAH ratios to apportion MGP tars according to their manufacturing process (Figure 7.7). The PAH indicators provide a vague classification of tars produced by coal carbonisation and CWG processes, however, there is a high degree of overlap between the regions which can incorporate a degree of subjectivity into the interpretation process. Therefore, a more robust multivariate approach was developed as part of this research to include a greater range of variables and thus improve specificity.
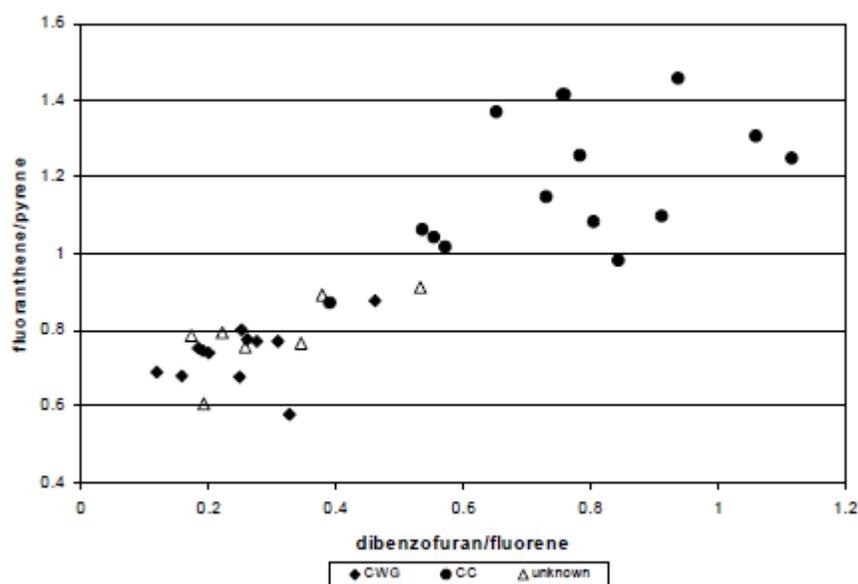
Figure 7.7: Diagnostic ratio plot for classification of coal tar type, where CWG=carbureted water gas tars and CC=coal carbonisation tars and unknown=samples of unknown origin (copied from Mauro, 2000).

### 7.1.4   Aim and Objectives

The overall aim of this Chapter was to provide a comprehensive evaluation of univariate and multivariate approaches for source apportionment of coal tars based on chemical fingerprints obtained by two-dimensional gas chromatography. To achieve this aim, the following objectives were established and met:

- The comparison of univariate and multivariate statistical analyses for coal tar source allocation.

- The investigation of a variety of data pre-processing methods to obtain optimal class separation.

- The use of historical site data to compare site manufacturing processes with source allocation.

## 7.2   Materials and Methods

### 7.2.1   Data Collection

Peaks were identified within each sample chromatogram using the GCxGC ChromaTOF® software (Leco, version 4.22). The data processing method computes the baseline, finds peaks, searches the library using the mass spectra of the identified peaks and finally computes the area/height of each identified peak. Certain parameters are entered to tailor the search for specific chromatograms. For example, the average peak widths in both dimensions, the minimum signal-to-noise ratio (S/N = 10 in this case) and the mass range can all be specified.

The full GCxGC chromatographic datasets of all coal tar samples were collated in a Microsoft Excel (version 12) spreadsheet using the Statistical Compare package within ChromaTOF® software (Leco, version 4.22). Statistical Compare aligns the chromatographic peaks identified within multiple samples. In total, the dataset consisted of 3479 variables (peak areas) but many of the peaks were only found in certain samples and not in every coal tar in the library.

For the reduced dataset of 156 selected peaks, the relevant peaks were identified (after data processing by ChromaTOF®) in each sample based on their retention times in the 1st and 2nd dimensions and by comparison of mass spectra to the reference library. The peak areas were manually collated in Microsoft Excel (version 11.8) prior to preprocessing transformations. The normalised reduced dataset of 156 peak areas is provided in Appendix C.

### 7.2.2   Data Collation and Preprocessing

Firstly, variables were normalised using the mass of coal tar extracted for each sample. The variables were then normalised to the internal standard by dividing the peak area of each compound by the peak area of the internal standard (D10-phenanthrene). Preprocessing strategies, such as various root, reciprocal and logarithm transformations were then performed within Excel prior to transfer of the dataset to a statistical software package.

### 7.2.3   Univariate Statistics

Univariate/bivariate plots and concentration graphs were compiled using a combination of Microsoft Excel (version 11.8) and Minitab® (version 16, Minitab Ltd., Coventry) software.

### 7.2.4   Multivariate Statistics

PCA, HCA and two-dimensional HCA heatmaps were all performed using Matlab® (R2011a, version 7.12, Mathworks Inc., Massachusetts). For HCA, only agglomerative clustering with Euclidean distance calculations were performed, but various linkage mechanisms (single, average and complete) were evaluated.

## 7.3   Results and Discussion

In this study, the focus of statistical data processing was to elucidate source-specific and/or process-specific differences between coal tar samples. Numerous methods of data interpretation were employed and are ordered in the following sections by increasing degree of classification ability, ranging from simple univariate methods to a number of different multivariate statistical techniques. The historical manufacturing processes employed at each site are restated in Table 7.1, with the addition of *site class* labels which will be used to identify group classifications within the statistical outputs throughout this Chapter.

### 7.3.1   Univariate Statistical Analysis

The first step in statistical analysis of the GCxGC coal tar dataset involved the comparison of samples using certain univariate/bivariate methods. The nature of such techniques is to focus on specific classes or groups of chemicals which may provide source differentiation. The main focus in coal tar samples is generally the PAH composition. The peak areas of individual components (as well as the PAH concentration data) were examined to determine whether any source-specific or process-specific trends could be discovered. Given the large volume of data, it is very difficult to identify trends within a table of numerical data, therefore, a selection of univariate statistical methods were

Table 7.1: Summary of manufacturing processes employed at FMGP sites investigated in this study with assigned site classes.

| DNAPL label | Site class[a] | Process(es) employed | Years of Operation |
| --- | --- | --- | --- |
| 1-6 | A | Vertical retorts; potential traces of CWG, oil reforming and early horizontal retort tar | 1836 - 1971 |
| 7 | B | Horizontal retorts | 1856 - 1969 |
| 8 | B | Horizontal retorts | 1886 - 1971 |
| 9 | A/E | Horizontal (early, low temperature) and vertical retorts plus CWG plant | ceased production by 1953 |
| 10 | B | Horizontal retorts | 1849 - 1981 |
| 11 | D | Wood preservation site - distillation of coal tar for creosote oil | unknown |
| 12 | E | CWG plant | 1854 - unknown |
| 13 | E | CWG plant | 1885 - unknown |
| 14 and 17 | B | Horizontal retorts; potential traces of CWG and coke oven tar | unknown |
| 15 | A | Tully gas plant; combination of vertical retorts and water gas | 1841 - 1961 |
| 16 | C | Horizontal retorts (early, low temperature) | 1854 - 1946 |
| 18 | F | Coke ovens | 1970's - present day |
| 19 | F | Coke ovens | 1930's - present day |
| 20[b] | A | Vertical retorts; potential traces of CWG tar | 1885 - unknown |
| 21 and 22 | C | Horizontal retorts (early, low temperature) | ceased production by 1870 |
| 23[b] | A | Vertical retorts | 1915 - 1979 |

[a] The probable site classes (A=vertical retort, B=horizontal retort, C=low temperature horizontal retort, D= creosote, E=carburetted water gas and F=coke oven) were assigned based on historical site data which indicated the periods of operation for each process. Class labels correspond to those given in multiple figures throughout Chapter 7.

[b] Sample used as part of a blind study, site details only disclosed (by Parsons Brinckerhoff) after analysis and data processing was completed.

employed. For example, a variety of graphical representations can be used to compare component concentrations and diagnostic ratios. The results of a number of these methods are summarised in the following sections.

## PAH distribution profiles

Quantification of the 16 priority pollutant PAHs present in all coal tar samples was performed as these compounds include the most hazardous chemicals present in coal tar. PAH quantification is generally one of the key stages in the investigation of contaminated sites such as FMGPs. The PAH contents of three petrogenic samples, petrol (P), diesel (D) and a 50:50 w/w mixture of petrol and diesel (PD), were also quantified to highlight the differences between petrogenic and pyrogenic distributions. The results of PAH quantification of coal tars and petrogenics are presented in Table 7.2.

The PAH concentrations varied greatly between samples, however, the same compounds were generally found in the greatest abundance; namely naphthalene and phenanthrene. It is difficult to visualise trends between the samples using this form of data representation, thus graphical representations were also employed. The major trends within the percentage composition of PAHs are shown in Figure 7.8. The samples are grouped by the major gas manufacturing method used at each site, presenting the average PAH composition for that particular process.

Table 7.2: US EPA priority PAH composition of all 25 coal tars and petrogenic reference samples.

| Sample | Concentration (mg/kg) | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | N | ACY | ACE | FLU | PHE | ANT | FLT | PYR | BaA | CHR | BbF | BkF | BaP | IP | DBA | BP |
| 1 | 16,797 | 4,206 | 883 | 1,917 | 3,341 | 3,511 | 1,773 | 2,079 | 688 | 710 | 45 | 306 | 3,282 | 42 | 196 | 272 |
| 2 | 9,867 | 3,891 | 407 | 1,334 | 2,730 | 1,308 | 860 | 1,414 | 441 | 388 | 79 | 174 | 1,046 | 83 | 48 | 62 |
| 3 | 4,617 | 1,605 | 244 | 774 | 1,743 | 945 | 962 | 1,074 | 372 | 326 | 77 | 170 | 899 | 43 | 23 | 55 |
| 4 | 5,788 | 1,857 | 252 | 832 | 1,930 | 984 | 975 | 1,076 | 416 | 348 | 117 | 226 | 917 | 79 | 29 | 78 |
| 5 | 1,068 | 503 | 93 | 267 | 567 | 437 | 269 | 292 | 103 | 59 | 11 | 45 | 363 | 13 | 15 | 32 |
| 6 | 481 | 252 | 57 | 122 | 271 | 235 | 223 | 216 | 84 | 46 | 16 | 32 | 267 | 13 | 14 | 26 |
| 7 | 47,171 | 8,309 | 1,365 | 5,612 | 21,304 | 11,298 | 16,589 | 14,526 | 6,763 | 6,674 | 3,187 | 4,855 | 15,367 | 2,303 | 606 | 2,263 |
| 8 | 85,432 | 16,567 | 1,485 | 6,970 | 29,445 | 7,194 | 13,630 | 11,821 | 5,184 | 4,759 | 1,957 | 6,408 | 18,823 | 2,133 | 913 | 2,183 |
| 9 | 31,763 | 5,044 | 1,455 | 3,644 | 10,892 | 7,180 | 9,220 | 7,165 | 3,265 | 3,706 | 1,656 | 2,543 | 11,506 | 1,010 | 490 | 1,271 |
| 10 | 81,931 | 7,131 | 1,538 | 5,251 | 20,107 | 8,982 | 14,988 | 12,844 | 5,538 | 5,307 | 2,990 | 4,362 | 15,016 | 2,165 | 600 | 2,168 |
| 11 | 36,169 | 2,333 | 22,458 | 15,764 | 36,896 | 15,630 | 10,518 | 7,926 | 2,228 | 2,447 | 663 | 828 | 3,137 | 306 | 62 | 371 |
| 12 | 4,222 | 924 | 6,678 | 6,877 | 11,361 | 7,774 | 8,157 | 6,141 | 4,904 | 4,426 | 2,472 | 2,894 | 3,371 | 1,395 | 868 | 1,332 |
| 13 | 10,152 | 10,400 | 9,418 | 14,149 | 26,529 | 10,344 | 10,770 | 12,657 | 3,112 | 2,651 | 759 | 746 | 1,935 | 1,082 | 95 | 680 |
| 14 | 8,255 | 13,035 | 3,158 | 12,683 | 22,879 | 8,310 | 14,898 | 15,238 | 3,906 | 3,569 | 1,583 | 1,498 | 3,158 | 1,665 | 278 | 1,593 |
| 15 | 79 | 97 | 16 | 78 | 284 | 115 | 182 | 206 | 82 | 80 | 42 | 33 | 87 | 109 | 2 | 73 |
| 16 | 3,111 | 2,392 | 247 | 1,679 | 4,668 | 2,023 | 2,773 | 2,329 | 1,174 | 890 | 488 | 477 | 791 | 533 | 50 | 337 |
| 17 | 4,341 | 7,096 | 1,670 | 5,758 | 17,615 | 6,836 | 17,130 | 12,816 | 3,964 | 3,713 | 2,018 | 1,855 | 3,855 | 2,164 | 217 | 1,888 |
| 18 | 7,148 | 12,120 | 727 | 8,408 | 30,056 | 14,677 | 22,509 | 27,013 | 9,342 | 10,296 | 5,992 | 5,165 | 8,620 | 4,647 | 615 | 4,632 |
| 19 | 7,031 | 10,332 | 1,570 | 14,909 | 36,448 | 17,232 | 29,459 | 32,731 | 11,703 | 13,292 | 8,283 | 7,453 | 11,358 | 4,655 | 1,663 | 5,646 |
| 20 t1 | 4,699 | 4,112 | 3,974 | 7,546 | 18,434 | 12,677 | 9,854 | 9,520 | 10,265 | 6,521 | 6,247 | 5,713 | 6,503 | 4,461 | 1,064 | 4,591 |
| 20 t3 | 5,154 | 3,997 | 4,992 | 6,943 | 17,205 | 11,597 | 9,244 | 9,286 | 9,947 | 5,954 | 5,748 | 4,548 | 6,054 | 4,138 | 1,014 | 3,963 |
| 20 t5 | 4,780 | 3,775 | 4,527 | 5,817 | 22,632 | 9,571 | 7,923 | 7,807 | 8,036 | 5,338 | 5,121 | 4,450 | 5,376 | 3,491 | 840 | 3,308 |
| 21 | 149 | 90 | 89 | 195 | 396 | 326 | 303 | 472 | 295 | 294 | 141 | 140 | 237 | 64 | 30 | 47 |
| 22 | 1,069 | 492 | 106 | 544 | 1,249 | 379 | 563 | 661 | 581 | 509 | 201 | 259 | 546 | 152 | 58 | 86 |
| 23 | 3,060 | 1,602 | 202 | 853 | 2,537 | 975 | 1,295 | 1,384 | 689 | 635 | 138 | 238 | 1,239 | 207 | 59 | 187 |
| P[a] | 835 | 12 | 7 | 22 | 26 | 5 | 1 | 8 | 13 | 5 | 1 | 2 | 61 | 1 | 2 | 5 |
| D[a] | 167 | 10 | 25 | 130 | 220 | 9 | 6 | 92 | 12 | 14 | 2 | 2 | 57 | 1 | 2 | 4 |
| PD[a] | 481 | 8 | 18 | 73 | 121 | 3 | 3 | 47 | 9 | 8 | 2 | 2 | 48 | 2 | 1 | 5 |

[a]P = petrol, D = diesel, PD = petrol/diesel mixture (50:50 w/w)
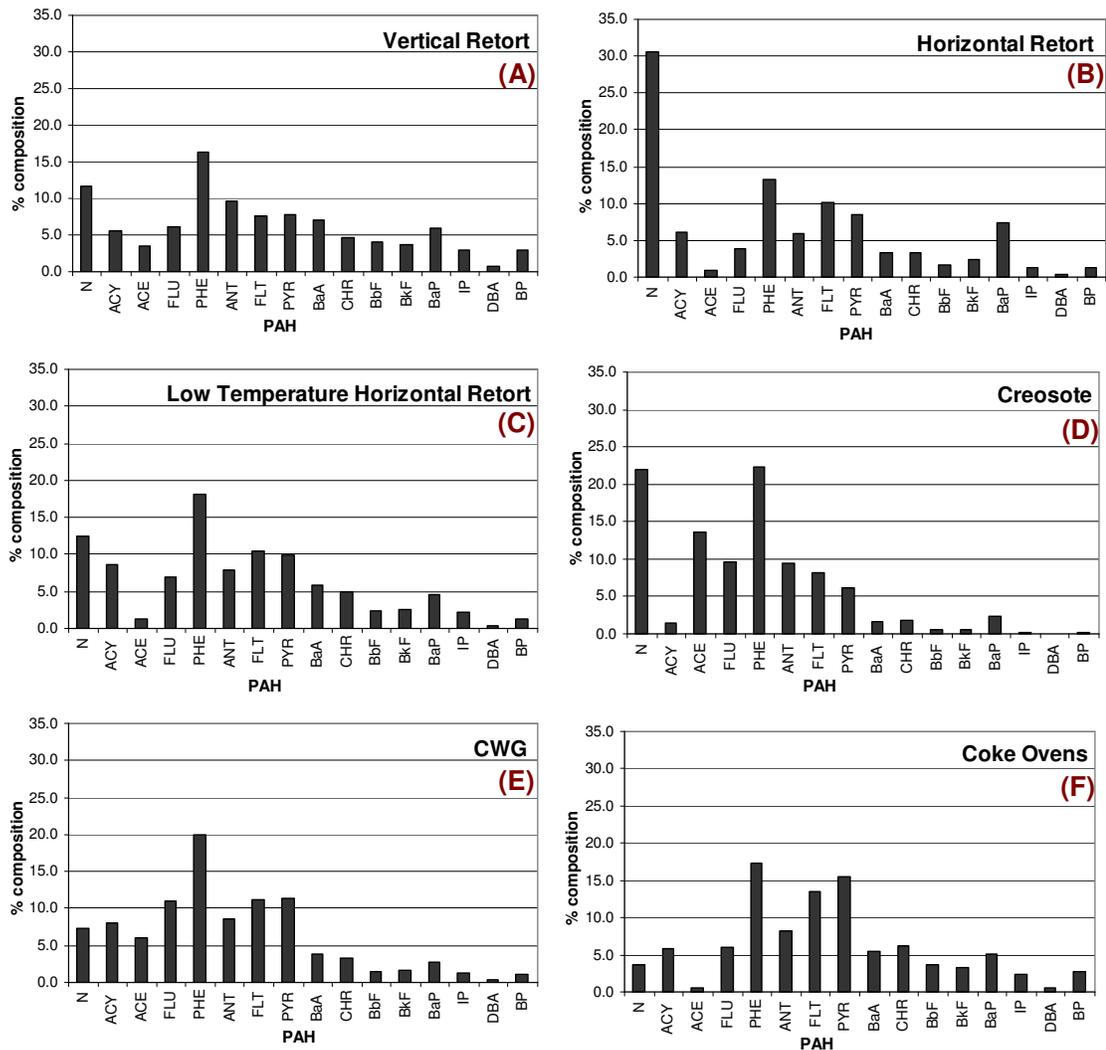
Figure 7.8: Graphs representing the percentage composition of each priority pollutant PAH (relative to the total priority pollutant PAH concentration) based on average values according to manufacturing process.

The vertical and low temperature horizontal retorts appear to exhibit very similar PAH distributions, most likely caused by the shape of vertical retorts providing a high proportion of tar with low temperature character due to limited coal/gas contact with the heated retort walls. A trend can also be seen within the high temperature horizontal retort samples, which exhibit large naphthalene concentrations in relation to other PAHs. These high concentrations can mask out other peaks in the chromatograms making it difficult to compare samples using traditional one-dimensional GC.

Figure 7.9 illustrates the difficulties in distinguishing tars based on their chromatographic pattern alone. At first glance, vertical and horizontal samples produce very similar chromatograms. These similarities between different tar types emphasise the need for robust statistical processing methods instead of pattern recognition which can be prone to subjective interpretation.
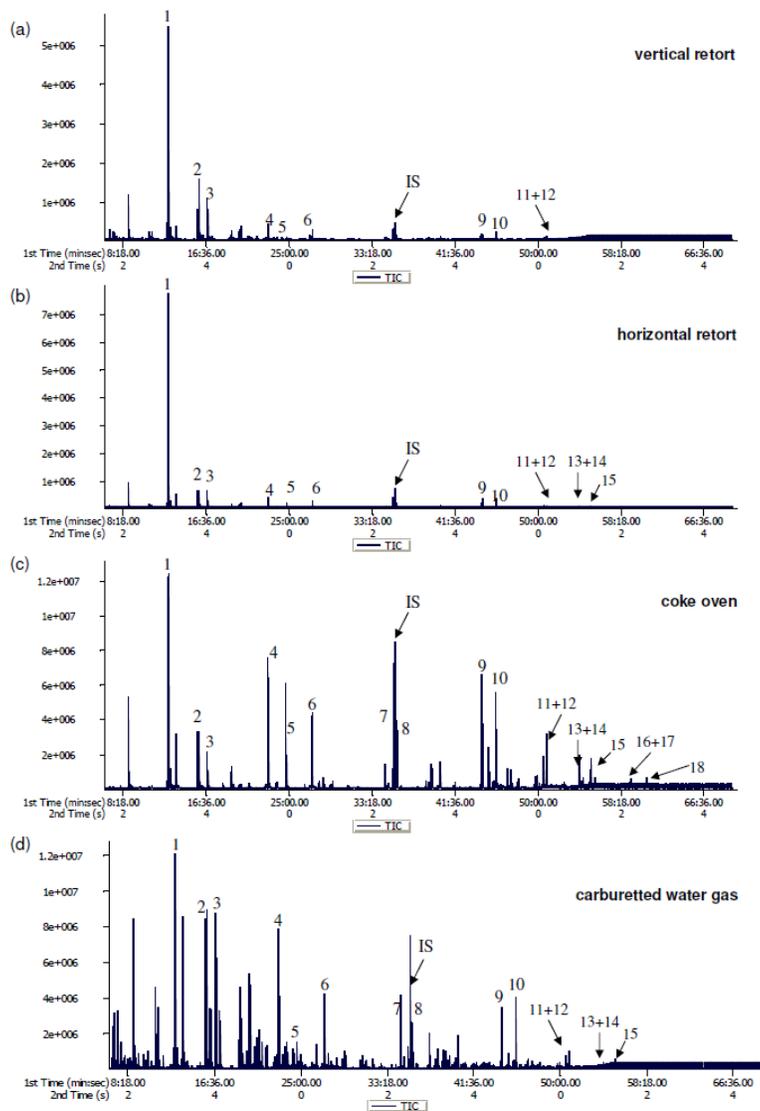


| #  | PAH | Abbr. |
|----|-----|-------|
| 1  | naphthalene | N |
| 2  | 2-methyl naphthalene | 2-MeN |
| 3  | 1-methyl naphthalene | 1-MeN |
| 4  | acenaphthylene | ACY |
| 5  | acenaphthene | ACE |
| 6  | fluorene | FLU |
| 7  | phenanthrene | PHE |
| 8  | anthracene | ANT |
| 9  | fluoranthene | FLT |
| 10 | pyrene | PYR |
| 11 | benzo[a]anthracene | BaA |
| 12 | chrysene | CHR |
| 13 | benzo[b]fluoranthene | BbF |
| 14 | benzo[k]fluoranthene | BkF |
| 15 | benzo[a]pyrene | BaP |
| 16 | indeno[1,2,3-cd]pyrene | IP |
| 17 | dibenzo[a,h]anthracene | DBA |
| 18 | benzo[g,h,i]perylene | BP |

Figure 7.9: Comparison of reversed phase GCxGC chromatograms for tars produced by various manufacturing processes. Numbering corresponds to the PAHs listed in the additional table. Chromatograms were formed by analysis of (a) sample 1, (b) sample 8, (c) sample 18 and (d) sample 14.

**Radial/Star plots**

An alternative method of comparing PAH concentrations and tar composition patterns employs radial (or star) plots. Radial plots allow any patterns in composition to be visualised in a simple manner, where each spoke represents a different variable. For example, the plots shown in Figure 7.10 clearly illustrate the similarities and differences in PAH composition between different coal tar samples. Figure 7.10(a) represents the pattern matching of three tar samples, 20t1, 20t3 and 20t5, obtained from the same tar plume at site S13, while Figure 7.10(d) depicts the major differences in PAH composition observed between tar samples from different types of gas plant (namely CWG, coke oven and vertical and horizontal retorts).

Figure 7.10(b) represents coal tar samples 1-6 from site S1 and demonstrates that the PAH components are found in the same ratios, but in differing concentrations across the site. There appears to be a significant difference in PAH ratio between fluoranthene (FLT) and pyrene (PYR) within sample 2. Values of this ratio are thought to be process-specific between CWG and retort carbonisation [Mauro, 2000, Saber et al., 2005]. Low FLT/PYR ratios are generally indicative of liquid fuel combustion while higher ratios (>1) are indicative of coal combustion. The FLT/PYR ratio obtained for sample 2 was 0.58, while the other samples from site S1 exhibited values in the range of 0.84-1.04. Therefore, it is likely that sample 2 contains a mixture of PAH sources, which is in agreement with the carbon isotope data for this sample which suggested a petrogenic origin. As site S1 also contained former CWG and reforming plants which used oil feedstocks, it is highly probable that sample 2 may contain a mixture of by-products from either of these processes combined with coal retort tar. Figure 7.10(c) also shows that the fluoranthene/pyrene (FLT/PYR) ratio is the main PAH compositional difference between samples 14 and 17 obtained from site S9. In sample 14, the calculated FLT/PYR ratio is 0.86, while in sample 17 it is 1.19. Site S9 was classified as mainly a horizontal retort gasworks, however, sample 14 was obtained close to a CWG tar emulsion tank while sample 17 was located nearby a large horizontal retort house. It would therefore be expected for sample 14 to have a lower FLT/PYR ratio indicative of petrogenic origin, due to traces of carburetion oil in CWG processes. This example shows that simple ratios can be a useful first step in distinguishing different contaminant origins.
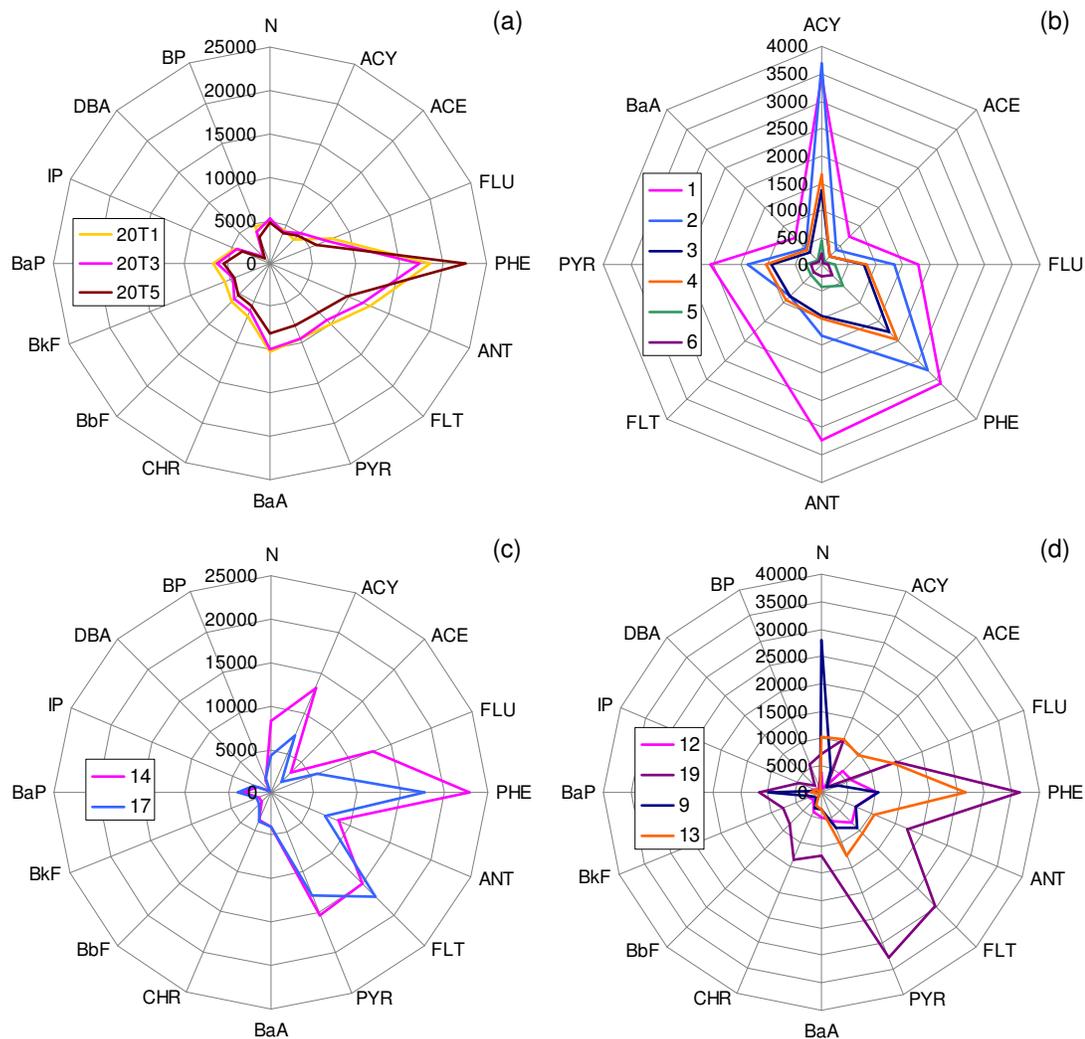
Figure 7.10: Radial plots indicating the PAH composition of coal tars from (a) site S13, (b) site S1 and (c) site S9 and (d) a comparison of tars from various different sources. Units of PAH concentration are shown in mg/kg and all PAH abbreviations are provided in Figure 7.9.

## Diagnostic Ratios

As previously mentioned, very few studies in the literature have compared coal tar chemical compositions for environmental forensic purposes. Mauro [2000] and Saber et al. [2005] demonstrated the use of PAH diagnostic ratios such as DBF/FLU and FLT/PYR for possible coal tar apportionment. Therefore, this was deemed the most logical place to begin investigation of diagnostic ratios.
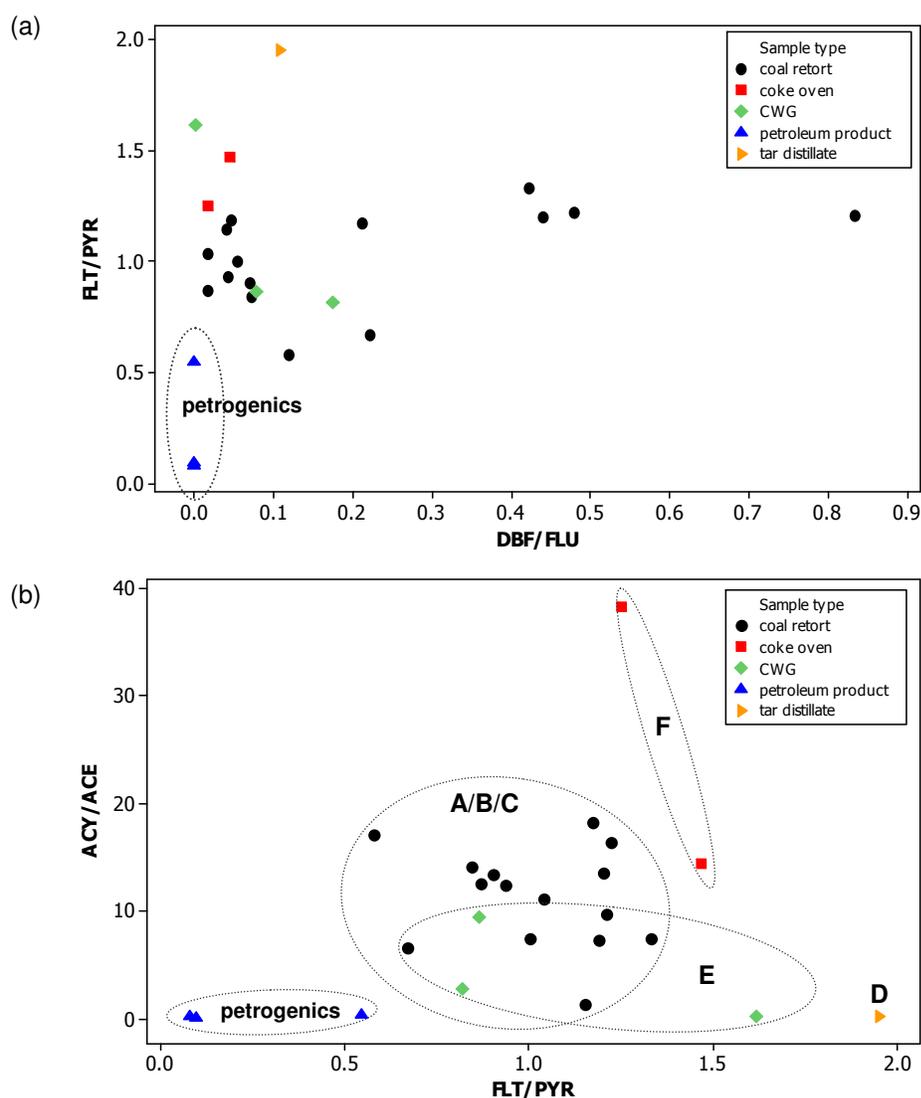
Figure 7.11: PAH double ratio plots for potential process-specific classification of coal tars.

Figure 7.11(a) illustrates the application of the double ratio plot suggested by Mauro [2000] to the GCxGC data obtained in this study. The diagnostic ratios were calculated using the normalised peak areas of the selected compounds (actual values of each ratio are provided in Appendix C). The results show very little classification of source, with the only real groupings between petrogenics and pyrogenics, yet overlap still exists with these classifications. In this research, the fluoranthene/pyrene ratio plotted against the acenaphthene/acenaphthylene ratio, as shown in Figure 7.11(b), was found to produce the best classification of all the simple comparison methods

investigated. Nevertheless, the plot is still unable to distinguish between retort tar sub-types. In addition, there is considerable overlap between CWG and retort tars, thus, univariate methods are not recommended for environmental forensic interpretation of coal tar DNAPLs. A multivariate approach is therefore deemed necessary for full differentiation of tar type.

## 7.3.2   Multivariate Statistical Analysis

Chemical analyses by GCxGC TOFMS produces extreme quantities of data on the composition of complex mixtures. The data can be described as a matrix of observed variables (peak areas) for each sample. The data matrices produced using GCxGC have high dimensionality; for example, a typical coal tar will produce over one thousand chromatographic peaks. Therefore, it is difficult to visualise the trends between multiple samples. Multivariate statistical analysis is an ideal method of reducing and simplifying GCxGC data to allow patterns between samples to be easily visualised.

A variety of multivariate statistical methods were investigated during the course of this study. The suitablity of hierarchical cluster analysis (HCA), two-dimensional HCA and principal component analysis (PCA) for describing the patterns and trends within the coal tar dataset are described in the following sections.

### Full chromatographic dataset

The full chromatographic datasets for all samples were compiled, as it was assumed that the more information included in the statistical model, the higher the likelihood of uncovering source-specific or process-specific trends. Prior to evaluation by multivariate statistical analysis, the peak areas of all components were normalised against the peak area of the internal standard (D10-phenanthrene) and subjected to a number of different preprocessing methods, including application of various root, logarithmic and reciprocal transformations. The preprocessing methods were used to standardise the dataset to prevent the differences between high concentration peaks from masking out the differences between those of low concentration. A variety of preprocessing methods were used to determine which strategy best described the dataset; a selection of PCA score plots produced using different data preprocessing strategies can be seen in Figure 7.12.
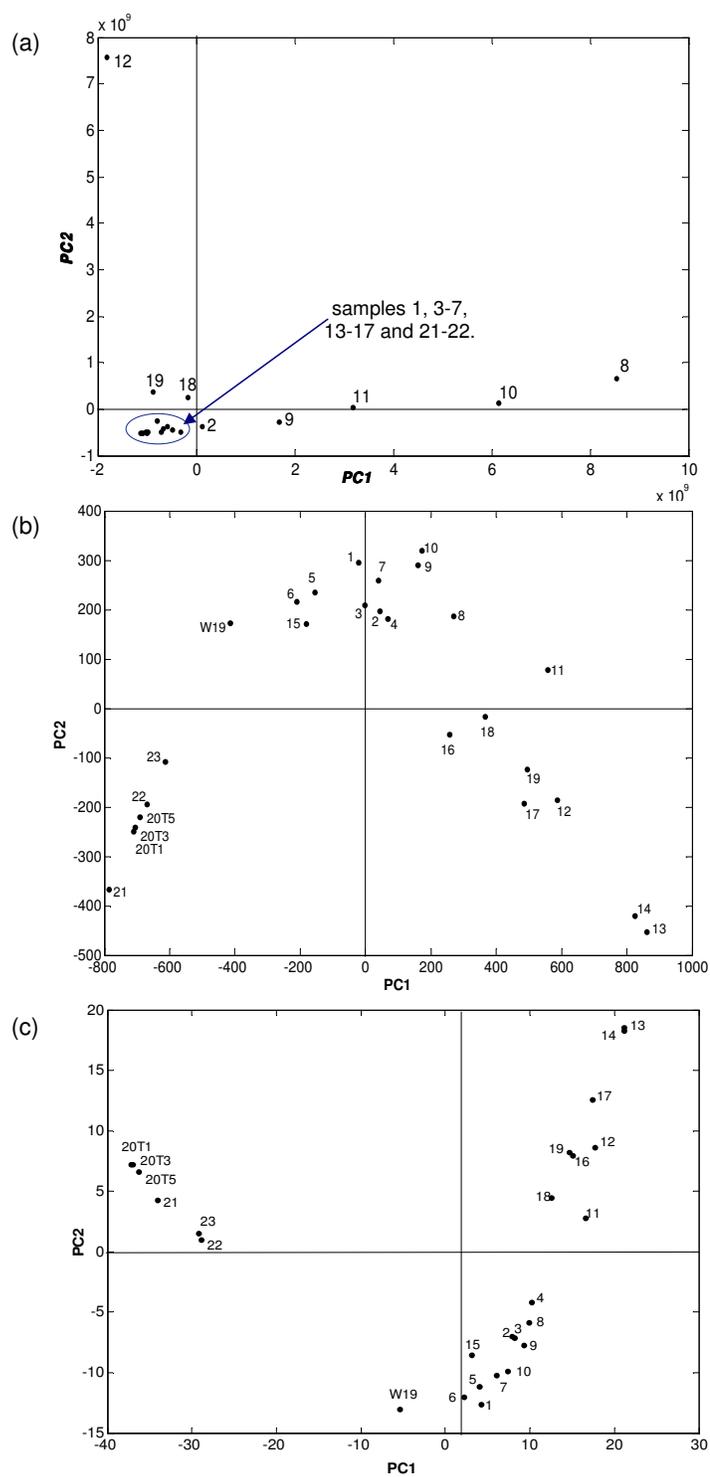
Figure 7.12: PCA score plots of the full chromatographic dataset using (a) normalised only, (b) normalised fourth root and (c) normalised sixteenth root transformations.

There are no clear clusters relating to source or manufacturing process in any of the PCA score plots produced using the full chromatographic dataset. Figure 7.12(a) illustrates that sample 12 is very dissimilar to the other tar samples. The large difference between sample 12 and the other tars is due to a far greater range of chemical components which masks the compositional similarities caused by manufacturing process. The application of various preprocessing strategies were not sufficient to allow discrimination of tar types (Figure 7.12).

**Reduced dataset**

The full dataset was not specific enough to provide source allocation of coal tars, therefore a reduced number of peaks which were sensitive to changes in manufacturing process were selected. The variables (peaks) were selected either based on their high loading scores (Figure 7.13) for the principal components produced in the score plots of the full chromatographic dataset (Figure 7.12) or their heightened intensity within at least one of the coal tar samples, thus likely to aid source differentiation. For example, it was discovered that n-alkanes were prevalent in CWG tars and were therefore included in the dataset. Chemical classes, such as phenols, which had been previously shown in literature to vary in concentration between different tar types were also included. In total, 156 peaks were ultimately selected, collated and pre-processed. A full list of the chosen peaks grouped according to chemical class and the normalised data used for statistical processing are provided in Appendix C. A summarised list of chemical classes is provided in Table 7.3 along with the corresponding class labels used to annotate certain statistical plots.
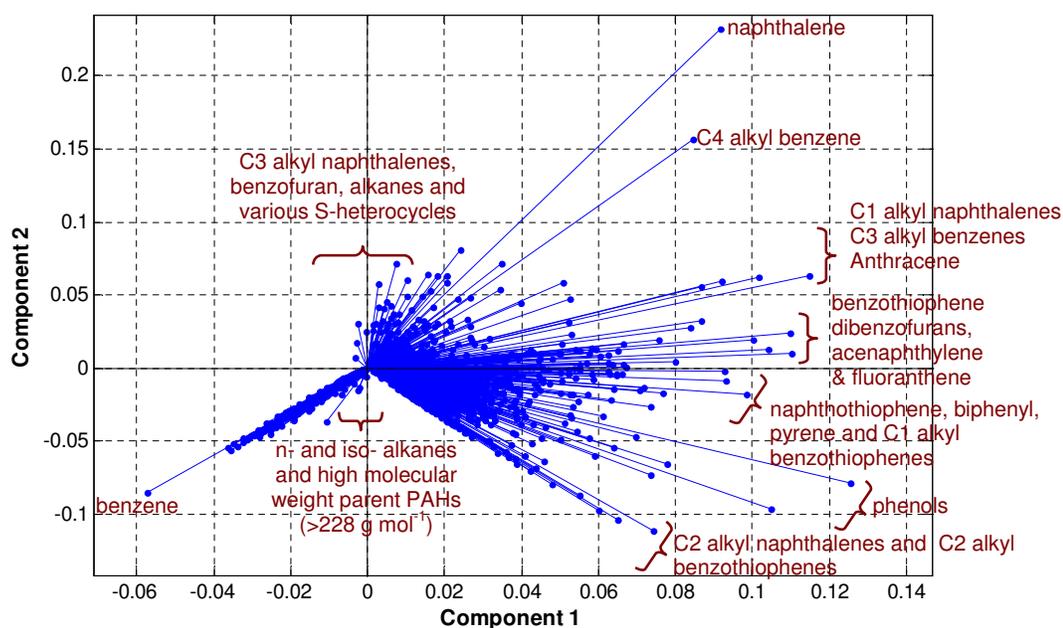
Figure 7.13: Loading plot for PCA of the full chromatographic dataset of 3479 peaks. High loading variables are annotated in red text.

Table 7.3: Summary of the chemical classes used within various statistical methods.

| Class no.[a] | Chemical class | No. of peaks used for each statistical method | | |
|---|---|---|---|---|
| | | HCA heatmap 1[b] | HCA heatmap 2[b] | PCA[b] |
| i | n-alkanes | 18 | 15 | 18 |
| ii | iso-alkanes | 5 | 5 | 5 |
| iii | alkyl benzenes | 11 | 11 | 11 |
| iv | phenols | 7 | 7 | 7 |
| v | hydronaphthalenes | 3 | 3 | 3 |
| vi | naphthalenes | 22 | 5 | 22 |
| vii | parent PAHs (≥3 rings) | 24 | 24 | 24 |
| viii | alkyl PAHs (≥3 rings) | 22 | 8 | 22 |
| ix | N-PAHs | 1 | 1 | 1 |
| x | O-PAHs | 5 | 5 | 5 |
| xi | parent S-PAHs | 5 | 5 | 5 |
| xii | alkyl S-PAHs | 34 | 10 | 34 |
| | Total | 156 | 99 | 156 |

[a] Classifications used within Figures 7.16 and 7.17.
[b]HCA heatmap 1, HCA heatmap 2 and PCA plot as given in Figures 7.16, 7.17 and 7.19 respectively.

**Hierarchical Cluster Analysis**

The dendrograms produced using a variety of data transformations and linkage methods were examined to determine which strategy best described the dataset. The ability of dendrogram linkages to describe the dataset was evaluated against the known sample provenance. For all data transformations, complete linkage was found to be the optimal linkage strategy. It was found that the "normalised only" dataset was not suitable for classification of the coal tars, while sequential root transformations aided the clustering (Figure 7.14). The dendrogram which best described the dataset was found using a normalised, eighth root data transformation. Successively higher root transformations were not able to differentiate between tar sources, it is thought that the transformations produced data too similar for effective linkages to be formed.
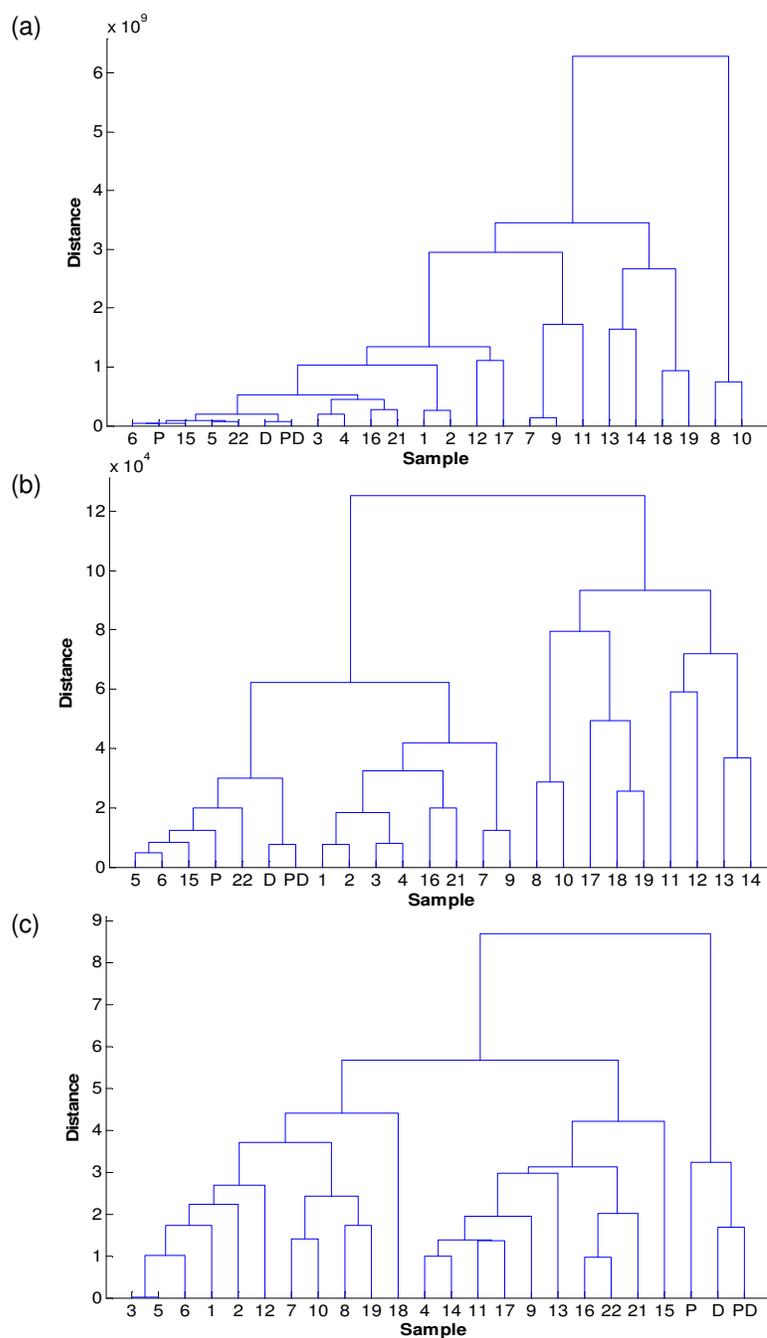
Figure 7.14: Dendrograms of the reduced GCxGC dataset using (a) normalised only, (b) normalised square root and (c) normalised reciprocal data transformations.

Figure 7.15 illustrates the HCA dendrogram which best described the dataset. Four main clusters were identified, however, within the cluster for horizontal retorts (labelled B/C) the coke oven tars

(F) were grouped separately. HCA was unable to distinguish between high and low temperature carbonisation processes, the low temperature retort samples were spread amongst the high temperature vertical and horizontal clusters. As previously mentioned, the high coal charge and small void space within vertical retorts typically produced tars of a low temperature character even though the process used temperatures in excess of 1000 °C to heat the retorts. It is therefore reasonable that samples 21 and 22 would group with the vertical retort samples as these samples were obtained from a former gasworks which is known to have only used early, low temperature processes. On the other hand, samples 16 and 9 were obtained from sites which likely used early, low temperature processes, but as it is difficult to get thorough historical data on gasworks sites, there is a possibility that the retort technology was upgraded. This could be the reason for samples 16 and 9 grouping with the high temperature horizontal retorts, they may contain a mixture of low and high temperature tar so have been classified accordingly.
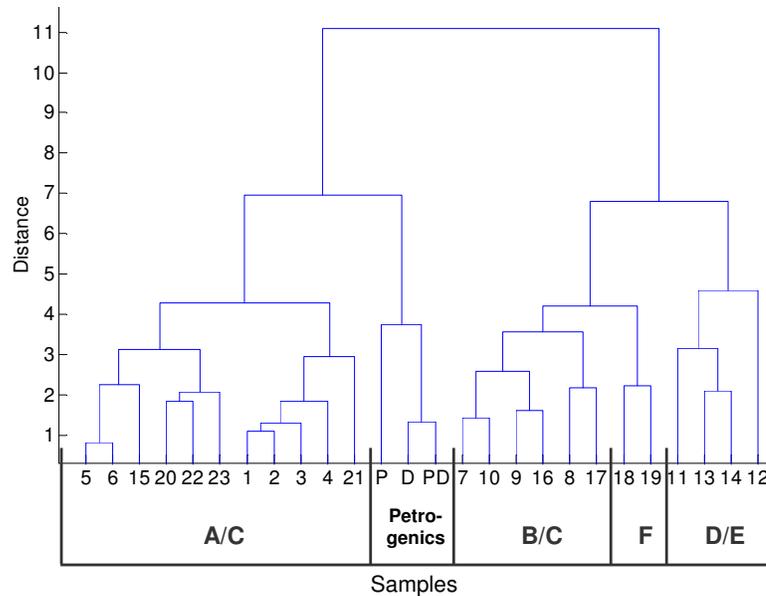


Figure 7.15: Dendrogram for the eighth root, normalised dataset using Euclidean distance and complete linkage mechanisms. Lettering corresponds to the sample provenance classifications given in Table 7.1.

Nevertheless, none of the tested data transformation and linkage combinations were able to fully differentiate between the tar samples based on manufacturing process. For this reason, alternative multivariate statistical methods were investigated.

**Two Dimensional Hierarchical Cluster Analysis**

As previouly described, a two-dimensional heatmap consists of two HCA dendrograms (to illustrate both sample and variable clustering) connected by a colour-coded matrix which reflects the intensity of each variable. In an attempt to discover the variables which caused sample clustering in the previous HCA plots, a two-dimensional HCA heatmap was prepared for the full dataset of 156 peak areas for 25 coal tars (however, due to their similar composition the peak areas of samples 20t1, t3 and t5 were combined and averaged for simplicity to give a single sample labelled as no. 20) and is shown in Figure 7.16.

Four main sample clusters were identified and the approximate groupings by manufacturing process are highlighted in blue text (based on the site classes listed in Table 7.1). Cluster D/E represents the CWG tars and creosote, however, the coke oven and retort tars do not fully separate and are spread out amongst the remaining three clusters. The second dendrogram associated with clustering of the variables (in rows) can provide extra information at a glance on the way in which the sample clusters have been formed. The shading of the heatmap mosaic indicates the differences between variables within each cluster. The chemical classes in Table 7.3 are identified within the heatmap by numbering next to the variable clusters. In an attempt to improve the classification power of the heatmap, the variables with poor differentiation between samples were removed. In general, alkyl-substituted PAH isomers showed similar responses in the heatmap and were deemed unnecessary for differentiation of the tar sources.

The variables included in this condensed dataset are indicated in Tables C.2 and C.3 of the Appendix by an asterix next to the included compounds. Where several isomers of a compounds were identified, an additional number in square brackets represents the number of isomers chosen for the condensed dataset. In total, 57 data points were removed and the HCA evaluation was repeated; the resulting heatmap is given in Figure 7.17. Once again, the number of peaks from each chemical class included in the heatmaps and the identity of chemical class labels can be found in Table 7.3.

The removal of unnecessary data points results in improved clustering for both samples and variables; as illustrated by Figure 7.17. Once again, four main sample clusters were identified and are annotated according to the site class labels given in Table 7.1. The CWG/creosote samples are again grouped within a single cluster (labelled D/E), however, the coke oven tars are now separated within cluster F. Clusters A and B represent mainly vertical and horizontal retort tars respectively.
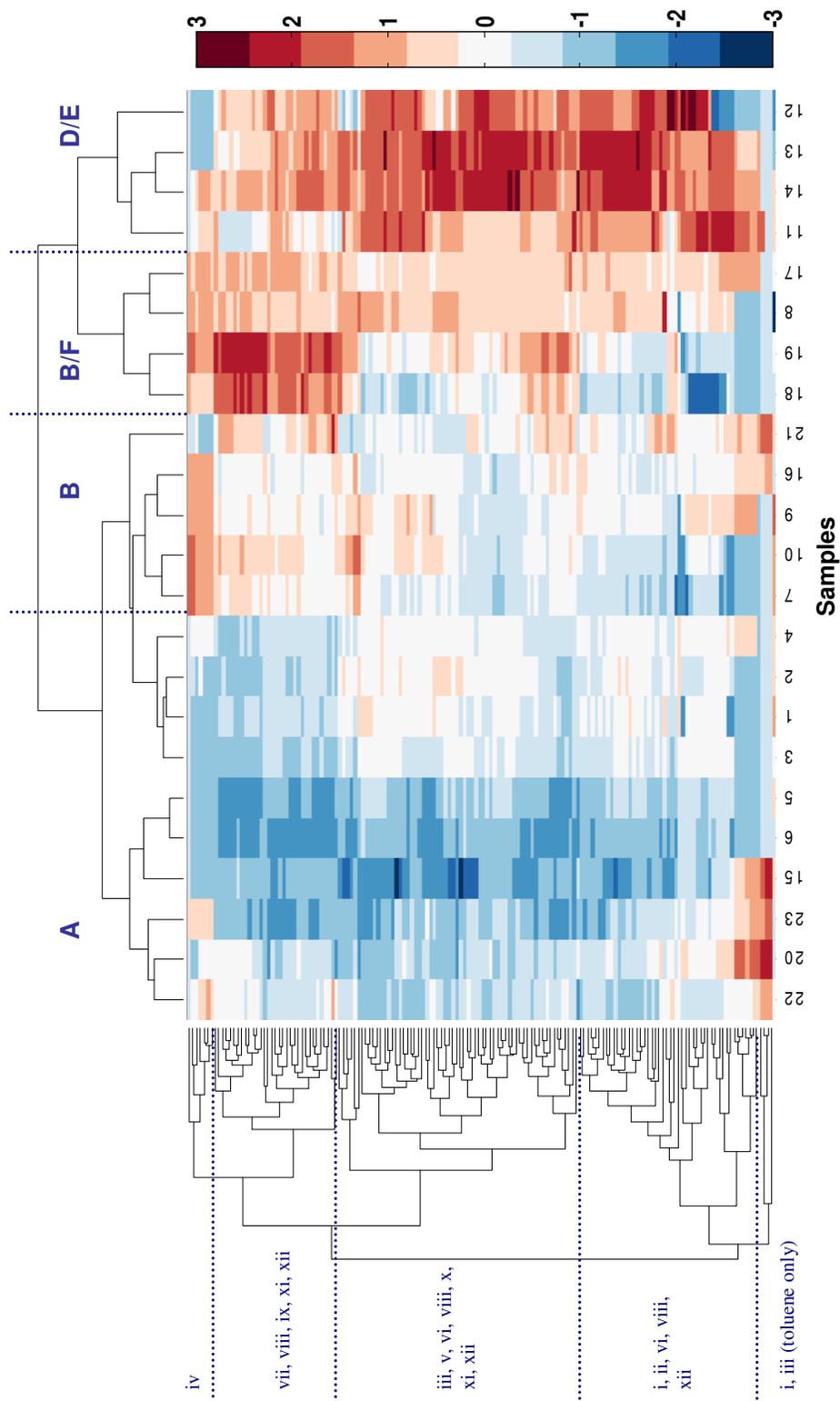
Figure 7.16: Heatmap of the full coal tar dataset. The red-blue colour gradient represents values of highest to lowest intensity [Lettering corresponds to site classes identified in Table 7.1, whilst numbering corresponds to chemical classes shown in Table C.2]. Reproduced from McGregor et al. [2011b].

However, samples 21 and 22, the low temperature horizontal retort (LTHR) tar, are still misclassi-
fied with the vertical retort tars. These rare samples were obtained from an FMGP which closed
in 1870. The early closure date means that the coal tars at this site could only have been produced
by low temperature horizontal retorts. As described in Chapter 2, the shape of vertical retorts
results in low contact of the majority of the coal charge with the hot retort walls. This results in
coal tar with a lower temperature character than that formed by horizontal retorts at equivalent
temperatures. Therefore, it is reasonable for the true LTHR tars to be classified as being more
similar to vertical retorts. As previously mentioned, samples 9 and 16 were also thought to be
produced by LTHR processes, however, it is possible that retort technology was upgraded at these
sites to give a mixture of tar sources. Furthermore, it is possible that environmental degradation
processes have altered the coal tar signature and interfered with the classifications as samples 21
and 22 were released into the environment over 140 years ago.

Sample 20 also exhibited unexpected clustering. The vertical retort tar was located within the
main retort cluster, however, it is represented by a single branch rather than as part of the vertical
retort group. The heatmap mosaic of the reduced dataset now shows defined sections resulting
in the process-specific clusters, allowing easy interpretation of the results. The shading intensity
represents the intensity of that variable, thus allowing the main differences in coal tar composition
to be found. The main chemical classes within each cluster are again labelled corresponding to
the numbering system for chemical classes shown in Table 7.3. Sample 20 and, to a lesser extent,
sample 15 have a high content of C27-C33 alkanes compared to the other retort tars. This deviation
in alkane content is sufficient to cause sample 20 to branch outside of the vertical retort cluster. It
is possible that these samples have resulted from multiple contamination sources (such as mixing
with an aliphatic-rich petrogenic source) or may simply have been exposed to less degradation than
the other retort tars studied. As previously mentioned, sample 20 (used here as an average of
the peak areas of sub-samples 20t1, 20t2 and 20t3) was obtained from a site which also hosted a
naphtha reforming plant. Any naphtha leakages are likely to have dissolved into the underlying
tar contamination at this site, thus altering the original coal tar chemical fingerprint. Naphtha has
a high aliphatic content so it is reasonable to assume that this may be the cause of the deviation
shown in the heatmap.

The main distinction between coke oven tars and other samples is the high parent PAH content
(Figure 7.17). The higher proportion of parent PAHs present within the coke oven tars is indicated

by the highly positive (dark red) shading for that cluster in the heatmap mosaic. The coke oven samples (18 and 19) were obtained on the day of production from an operational steel works, thus, high levels of parent PAHs were anticipated due to limited opportunity for degradation. Parent PAHs degrade faster than their alkylated homologues, therefore, weathered pyrogenic samples generally have a characteristic PAH pattern of $C0 < C1 < C2 < C3$, while parent PAHs are dominant in fresh pyrogenic samples.

The heatmap also depicts the main variation between horizontal and vertical retort tars. Initially, using the full dataset, the two retort types could not be differentiated (Figure 7.16). With the condensed dataset, the retort types are now clearly separated into two well-defined clusters, with the exception of low temperature retort samples 21 and 22 (Figure 7.17). The heatmap shading indicates that the retorts produce different quantities of phenol/alkyl phenol compounds, and as such, cluster analysis separates horizontal retorts (that generally have a high content of phenols) from the vertical retorts (that have little or no phenols present). This variation is likely a result of the length of time in which hot gases evolving in the retort are kept in contact with the hot retort walls. In horizontal retorts, the gaseous compounds have a greater opportunity for further degradation and higher degree of oxygen and water vapour availability due to the greater void space, thus explaining the presence of phenols. Recent work by Richards et al. [2011] as part of a MRes project at the University of Strathclyde, investigates a one-step method for extraction, derivatisation and GCxGC TOFMS analysis of coal tars. The method further elucidates coal tar composition by providing derivatisation of labile-H containing compounds, allowing phenolic content to be studied with greater accuracy. The derivatisation method could also prove particularly useful in discriminating between vertical and horizontal tars, since vertical tars are thought to contain high levels of tar acids which may not be identified without derivatisation. Furthermore, the technique could elucidate possible biodegradation products and aid natural attenuation studies at FMGP sites.

Distinct differences are shown in Figure 7.17 between the CWG and coal carbonisation tars. The CWG tars exhibit a far greater concentration of low molecular weight alkanes, alkylated benzenes and sulfur-containing heterocycles than other tars. The additional oil spray used within the gas enrichment step at CWG plants could easily account for the high levels of low molecular weight aliphatics and aromatics in these tars, as they are all found in high abundance in petroleum products. The enhanced levels of sulfur-containing heterocyclic PAHs, such as benzothiophenes and

dibenzothiophenes, may also be explained by the addition of oil. Heavy oil or lighter petroleum products could be used in the CWG process and were generally chosen based on availability and cost; therefore the sulfur content may be a useful way of distinguishing different CWG sources. On the other hand, the high S-PAH content of CWG tars may be due to the use of lower quality, inexpensive coals (with higher sulfur contents) for the formation of the coke used within CWG systems.

## Principal Component Analysis

As HCA and 2D HCA were unable to fully differentiate between coal tar classes, principal component analysis was investigated. A fourth root data transformation was found to provide PCA score plots which best described the dataset. Without data preprocessing, the large range of peak intensities within the dataset results in small peaks contributing less towards the principal components, regardless of their chemical importance. This is due to the strong link between the mean value of an analyte and it's variance. Trace analytes give lower variance than analytes which are found in high concentrations. Moreover, the samples have been residing in the subsurface environment for decades, if not centuries, therefore environmental processes may have changed the contaminant ratios due to evaporation or dilution into groundwater. Taking the fourth root of the data allows the focus to be on the presence/absence of contaminants rather than their concentrations, and as such relates to the primary production method instead of changes which may have occurred due to environmental factors.

The PCA technique was first attempted using a dataset composed of the 16 EPA PAHs and their alkyl derivatives (grouped in bands according to alkylation level) which are currently the contaminants typically measured during investigations at former MGP sites. The resulting PCA score plot (Figure 7.18) provided loose clustering of samples 1-6 from site S1 as well as the coke oven tars (samples 18 and 19), however source-specific and process-specific allocation was not possible. This emphasises the gap in scientific knowledge of coal tar composition which ultra-resolution chemical fingerprinting can resolve.

194
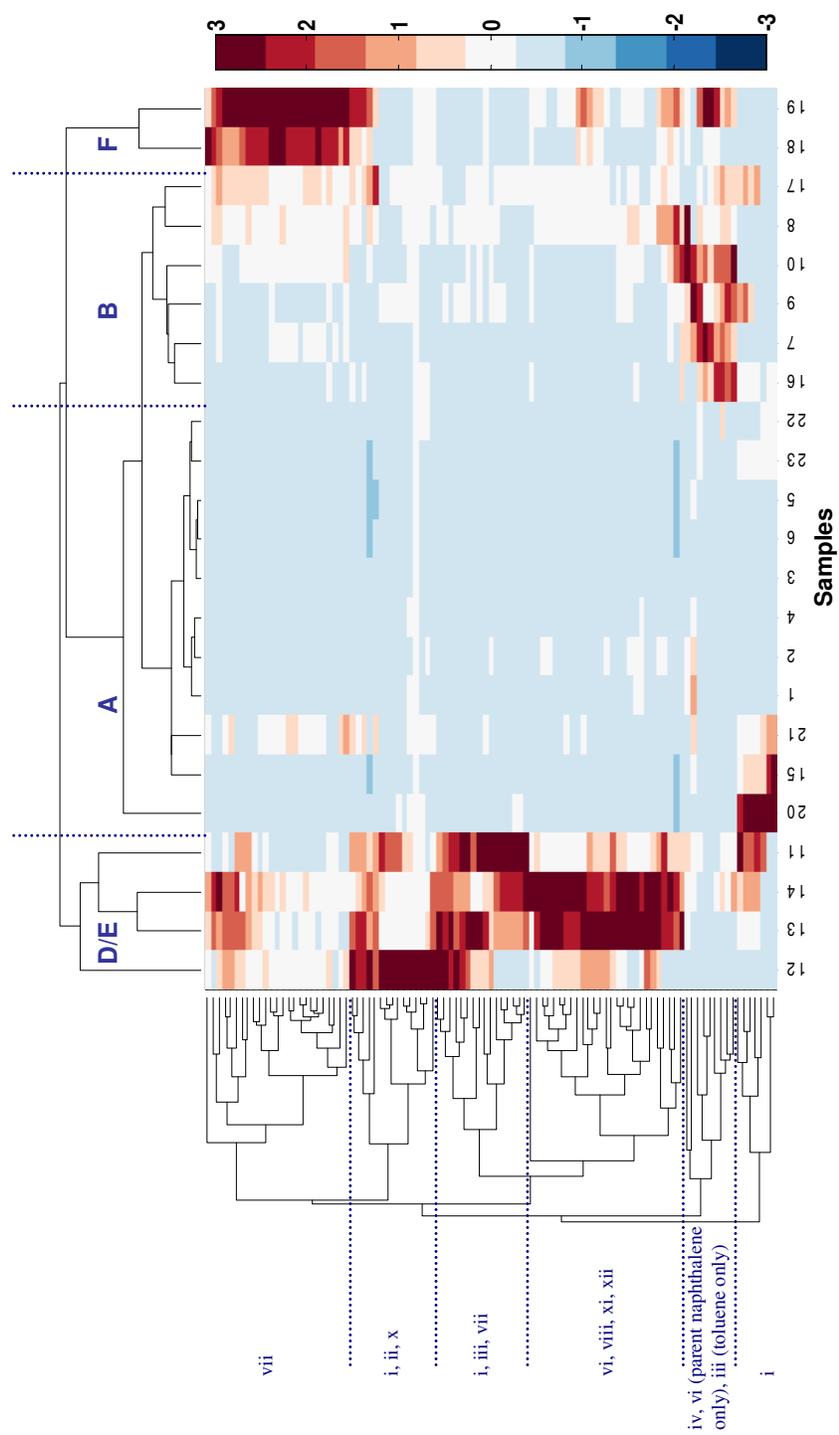
Figure 7.17: Heatmap of a reduced coal tar dataset. The red-blue colour gradient represents values of highest to lowest intensity [Lettering corresponds to site classes identified in Table 7.1, whilst numbering corresponds to chemical classes shown in Table C.2]. Reproduced from McGregor et al. [2011b].
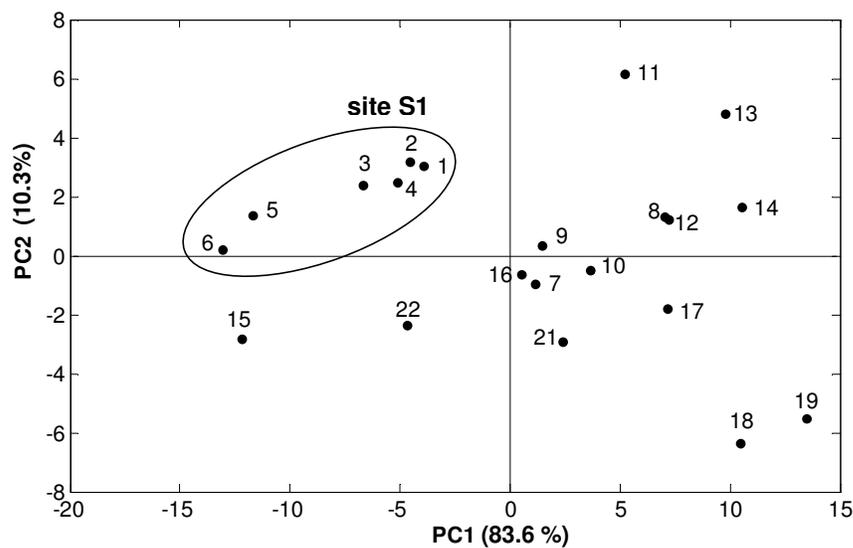
Figure 7.18: PCA score plot of 16 EPA PAHs and alkyl homologues (grouped by alkylation level) using the coal tar GCxGC dataset.

The GCxGC dataset of 156 peaks was then subjected to principal component analysis. The resulting PCA scoreplot of the normalised fourth root dataset is given in Figure 7.19. The first two principal components (PCs) describe 83% of the total variance within the dataset. Arbitrary boundaries have been constructed to highlight the groupings achieved; once again the lettering corresponds to site classes as given in Table 7.1. Unlike Figure 7.18, the class boundaries in Figure 7.19 show well-defined separation of coal tars based on manufacturing processes used in their production. The retort tars can now be separated into their three sub-types; vertical (A), horizontal (B) and low temperature horizontal (C) retorts. The additional cluster of low temperature horizontal retort comprises of samples 9, 16, 21 and 22. Based on historical data of site processes, it was anticipated that tar samples 16, 21 and 22 (class C) were produced by low temperature horizontal retorts, however, sample 9 was expected to have been influenced by later vertical retorts and CWG processes on site. This alone indicates how powerful the ultra-resolution method can be for source apportionment of coal tar plumes.

The creosote oil sample is again grouped within the CWG cluster, as in both HCA heatmaps. There are two possible explanations for this; (1) the tar used to produce the creosote oil via distillation was produced by a CWG plant, or (2) the high aliphatic content of the medium distillate creosote oil is simply most similar to that of CWG tar and is thus grouped accordingly. The creosote tar

clusters closely with sample 12; a tar acquired from a complex FMGP where on-site distillation of tar was most likely performed, yet the abundance of higher molecular weight PAHs in sample 12 suggest it is not a product of distillation. The inclusion of a greater number of creosote samples in the score plot is required to fully understand the reasons for this grouping.
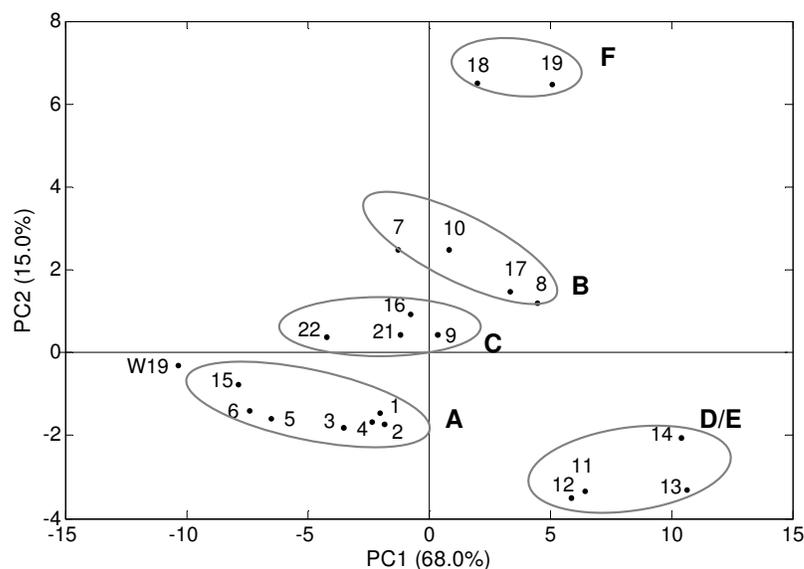


Figure 7.19: PCA score plot of the initial sample set with a normalised, fourth root data transformation; lettering corresponds to site classes as identified in Table 7.1 [where A=vertical retort, B=horizontal retort, C=low temperature horizontal retort, D= creosote, E=carburetted water gas and F=coke oven].

It can be seen that the high and low temperature horizontal retort tars (classes B and C respectively) cluster in a similar manner across the upper quadrants of the score plot, however, appear to be separated out into two closely clustered groups by PC2. This correlates well with the variable loadings provided in Figure 7.20. PC2 shows highly positive loadings for the phenols, which have previously been shown to vary in quantity between different retort tars [Young, 1922, Malatesta, 1920, Terrace, 1948]. The loadings were further investigated, showing that phenols, many high molecular weight parent PAHs and alkanes (C24-C29) were the most relevant variables for PC2 and were thus the peak areas with most variation between the retort tar sub-types and the coke oven tars. This further correlates with the expected differences between the fresh coke oven tars (class F) and the historic retort/CWG tars, as the lower incidence of weathering allows for higher levels of parent PAHs in the fresh tars.

PC1 was mainly defined by acenaphthene, sulfur heterocycles, alkyl benzenes and C1 methyl naphthalenes (highly positive loadings) and C30-C33 alkanes (negative loadings), and mainly explained the differences between CWG/creosote tars and the other tar types. Once again, these loadings fit well with the observed clustering of samples in the score plot. The heightened sulfur content of CWG tars is evident in their higher PC1 scores than other tar types. It was also noted, through examination of GCxGC chromatograms, that the CWG tars exhibited an increased content of alkyl benzenes relative to the other tars.
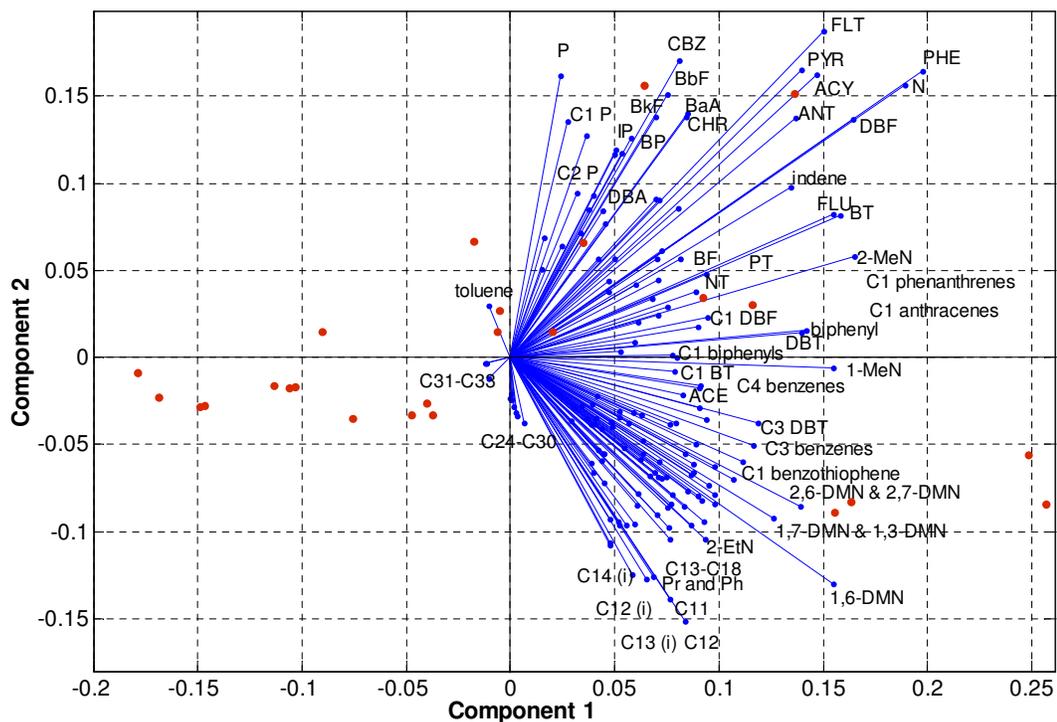


Figure 7.20: Loadings plot corresponding to the PCA score plot provided in Figure 7.19. Key : N=naphthalene, P=phenol, ACE=acenaphthene, ACY=acenaphthylene, FLU=fluorine, PHE=phenanthrene, ANT=anthracene, FLT=fluoranthene, CHR=chrysene, BbF=benzo[b]fluoranthene, BkF=benzo[k]fluoranthene, DBA=dibenzo[a,h]anthracene, BP=benzo[g,h,i]perylene, IP=indeno[1,2,3-cd]pyrene, BF=benzofuran, DBF=dibenzofuran, MeN=methylnaphthalene, DMN=dimethylnaphthalene, PT=phenalenothiophene, NT=naphthothiophene, BT=benzothiophene and DBT=dibenzothiophene.

The basic site details for all coal tar samples are summarised in Table 7.1, where it can be seen that samples 14 and 17 were obtained from the same FMGP. The samples were obtained from different sampling locations within the site, yet as Figure 7.19 illustrates, they are not clustered

closely in the PCA score plot. Sample 14 clusters with the CWG tars, while sample 17 is clearly grouped with the horizontal retort tars. Historical site data shows that sampling site of 14 was located nearby a CWG structure, whereas the location of sample 17 was close to the horizontal retort house. This demonstrates that GCxGC analysis coupled with PCA has the capability to differentiate between tar signatures, not only from different sites, but between multiple coal tar DNAPL plumes across a single site also. The technique has the potential to allocate multiple tar plumes to the specific time period in which a particular manufacturing process was in operation. The high degree of contamination present at most FMGP sites results in such sites being blamed immediately for any PAH contamination found in the vicinity. The level of knowledge on coal tar composition obtainable by GCxGC with PCA has the potential to easily settle any debates over liability at FMGP sites and the surrounding area.

The PCA model allows classification of five tar types investigated in this study, with the potential for site-specific differences within each cluster to be identified. The model could prove invaluable for source allocation of FMGP wastes, by identifying the specific process(es) used to produce tar plumes across a site, and thus the operator(s) responsible for the contamination. An additional PCA score plot provided in Figure 7.21 represents the potential of the model to differentiate between multiple sources of contamination. Figure 7.21 combines the initial coal tar library with a set of three petrogenic samples (petrol, diesel and a petrol/diesel mixture) which were extracted and analysed in the same way as the coal tar samples. The petrogenic samples can be seen to cluster far from the pyrogenic coal tars, indicating that the PCA model is not only capable of differentiating between coal tars but can be used to discover additional sources of contamination across a site which were not produced by the gasworks processes. The potential of the PCA model for classifying a greater variety of sources could be evaluated by expansion of the dataset. Another interesting feature in Figure 7.21 is the clustering of samples 1-4 (site S1) in the same quadrant of the score plot as the petrogenics, indicating similarities in character. This reinforces the results obtained by the carbon isotope studies provided in Chapter 5 which indicated that samples 2 and 3 had carbon isotope signatures in a region most associated with petrogenic contaminants. It is thought that mixing of pyrogenics and petrogenics may have occurred on this site due to the presence of an underground storage tank for heavy oil.

The model could also be further validated by expansion of the dataset and may potentially allow additional gas manufacturing processes to be classified, especially if worldwide samples are obtained.

For example, tars produced by a number of additional MGP processes (such as producer gas and Mond gas plants) could provide further clusters in the score plots. In addition, tars produced by feedstocks other than coal are expected to provide distinctly different chemical fingerprints, which could give interesting groupings when input into the PCA model. For example, it was not uncommon for wood or heavy oil to be used as the primary fuel source at FMGPs in certain parts of the world when coal supply was limited. Therefore, PCA score plots of a range of global FMGP tar samples could potentially indicate differences in the raw materials used in gas production.
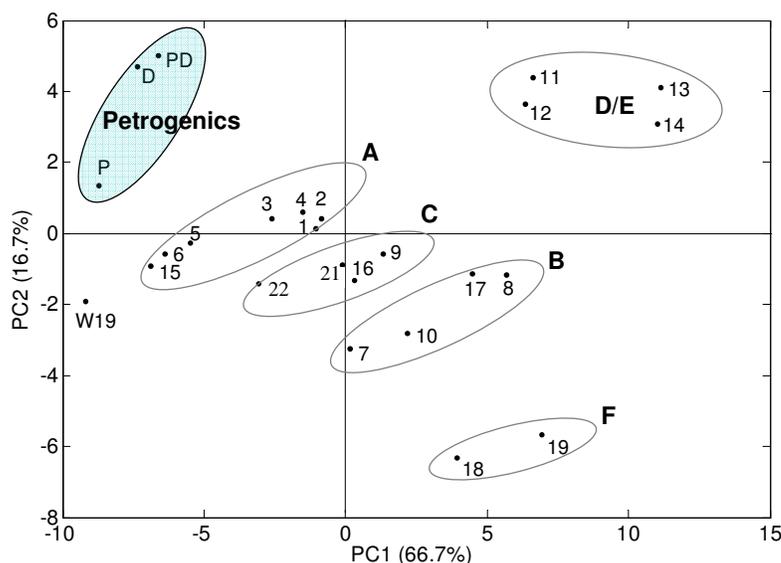


Figure 7.21: PCA score plot of the initial sample set with a normalised, fourth root data transformation including a set of petrogenic samples (P=petrol,D=diesel and PD=petrol/diesel mix (50:50 v/v); group classifications correspond to site classes as identified in Table 7.1.

**Weathering**

As previously stated, the coal tar samples analysed in this study have been residing in the subsurface environment for decades. Thus, it is likely that the samples have been subjected to a degree of weathering. The effect of weathering on coal tar composition is an area of research which could be greatly expanded upon, yet was outwith the scope of this thesis and as such is only investigated briefly here. The PCA plots previously described in this Chapter were obtained using a transformed dataset in an attempt to remove the effect of weathering and focus on source-specific chemical variation.
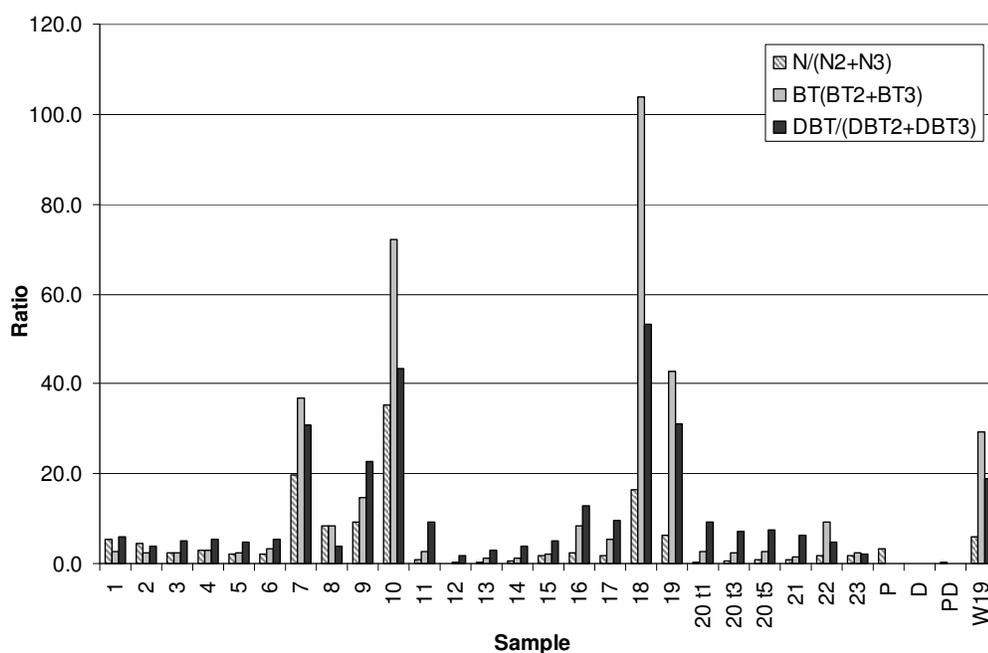
Figure 7.22: Comparison of weathering ratios within the coal tar library. N=naphthalene, BT=benzothiophene and DBT=dibenzothiophene.

The ability of PCA to estimate the degree of weathering in coal tars was evaluated by compiling a score plot for only the compounds most susceptible to weathering in addition to a number of established weathering ratios. The values for three PAH weathering ratios were included due to the tendency for alkyl homologues to be more prevalent in severely weathered samples. For example, a naphthalenes ratio comparing the abundance of parent naphthalene (N0) to C2 and C3 alkyl derivatives (N2 and N3 respectively) was calculated using the equation N0/(N2+N3). Equivalent ratios for benzothiophene (BT), dibenzothiophene (DBT) and their alkylated homologues were also calculated. Higher ratios represent samples with the least weathered character. The weathering ratios calculated for each coal tar extract are compared graphically in Figure 7.22 (the exact value of each ratio can be found in Table C.14 of the Appendix).

Figure 7.22 shows that samples 7, 10, 18 and 19 (and to a lesser extent sample 9) produce far greater ratios than the remaining tars. It was anticipated that coke oven tars 18 and 19 would produce high values as these samples were removed from coke ovens on the day of production, thus should not have experienced any effects of weathering. However, samples 7, 9 and 10 were all sampled at FMGP sites which have not been operational in at least 30 years. These particular samples were

located within former tar tanks and would therefore have been sheltered, to a certain extent, from weathering which may explain the higher values.

The values presented in Figure 7.22 were combined in a dataset with the peak areas of a number of individual compounds which are expected to be most susceptible to weathering. The few individual compounds included in the plot were the phenols, toluene, indane, indene and the hydronaphthalenes. The resulting PCA score plot for the weathering dataset shows great variation amongst the samples (Figure 7.23). An aliquot of fresh coke oven tar (sample 19) which had been subjected to three months of air evaporation (sample W19) was used as a reference marker, allowing a general trend in weathering across the score plot to be noted. The arrow in Figure 7.23 indicates an approximate trend from regions of low weathering (including fresh tar samples 18 and 19) to regions subjected to increasing degrees of degradation (such as sample W19). The air evaporation test (which produced tar W19) is not an accurate approximation of how coal tar would be degraded, yet provides a reference marker in this case. An ideal experiment to imitate weathering would involve water-washing of a fresh coal tar to discover the rate of dissolution and the compositional changes which may occur. In the original PCA plot shown in Figure 7.19, the weathered version of sample 19 clusters nearby the vertical retort tars, especially sample 15. It was expected that sample 15 had been exposed to significant weathering due to the sampling location on site S10, however, it cannot be firmly concluded that this is the reason for the close clustering of sample W19. Further investigation into the compositional changes in coal tar caused by weathering is required.

As expected, the petrogenic samples (P, D and PD) cluster close to the weathered sample (W19) in Figure 7.23 due to the reversal in PAH distributions found between petrogenic and pyrogenic sources. It is therefore difficult to determine whether tar samples cluster closely with the pyrogenics due to extreme weathering or because of a degree of petrogenic character in the tars. This research could be expanded in the future by the development of a mixing model for coal tars and common petrogenics alongside a water-washing degradation model.
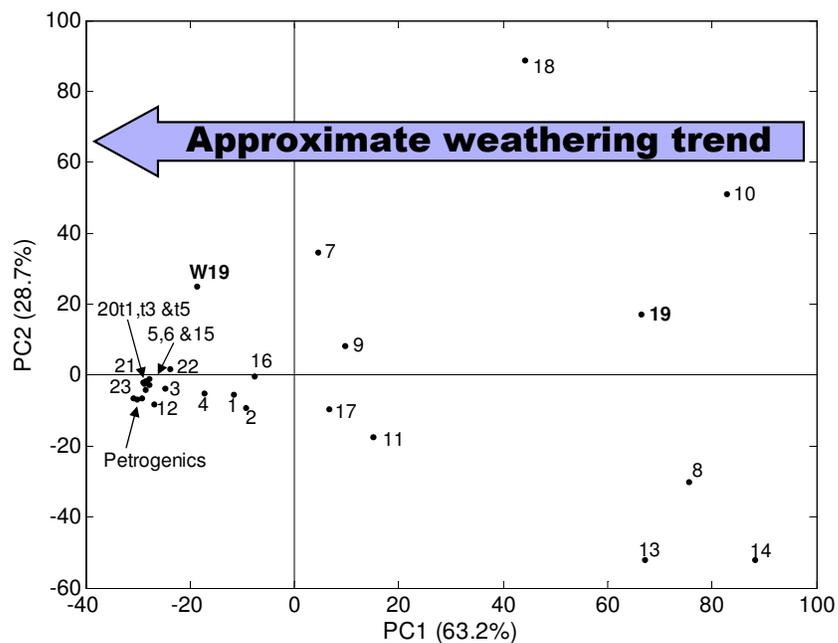
Figure 7.23: PCA score plot to compare the extent of weathering in coal tars; the arrow indicates approximate regions of low to high weathering.

## 7.4   Conclusion

The statistical methods shown in this Chapter have demonstrated that it is possible to classify coal tars based on the processes by which they were formed over 100 years ago. The enhanced analytical power of GCxGC TOFMS allows more chemical information to be gained per sample than conventional analytical methods, allowing for a more robust classification scheme. Multivariate statistical processing is required for the classification of coal tars, as univariate methods often exclude vital portions of data which provide source discrimination.

Preprocessing was shown to be an essential step in the statistical interpretation of coal tar data. The large range of peak intensities results in small peaks contributing less towards the statistical model unless data transformation is employed. Root transformations were shown to provide the optimal data clustering in both hierarchical clustering and principal component analysis.

PCA was shown to have the highest discriminatory power of all the statistical methods investigated

in this study, and as such a PCA model was developed capable of process-specific apportionment of coal tars. The following Chapter will test the validity of the PCA model through classification of additional coal tars as part of a blind trial.

# Chapter 8

# Ultra Resolution Chemical Fingerprinting as a Tool for Environmental Forensics

## 8.1  Introduction

The introduction of recent environmental legislation, such as the European Environmental Liability Directive (2004/35/EC) and Part IIA of the Environment Protection Act, has increased the scope and abundance of environmental forensic investigations in the United Kingdom. Environmental forensics is an established industry in the United States and Canada [CERCLA, 1980] due to the multitude of oil spill incidents which spurred its development, however it is still gaining purchase in the UK.

The landmark Corby case [Corby Group Litigation v Corby District Council, 2009] brought much attention to the use of environmental forensics in the United Kingdom. The town of Corby in Northamptonshire was affected with an unusually high rate of birth defects with respect to the UK average. The heightened defect rate was eventually linked to the remediation process at a former steel works between 1985 and 1999 [Corby Group Litigation v Corby District Council,

2009]. The local council was found negligent in the way in which they performed the remediation. Contamination from the site was shown to have spread to the local area during transportation to landfill causing residents to inhale and ingest dust containing chromium, cadmium, nickel, dioxins and PAHs.

With the high degree of contamination expected at FMGPs, and the sheer number of such sites in the UK (estimated at 3000±1000), it is reasonable to assume that they will be involved in many liability cases in the future. One such case has already been heard in the UK, involving a former gasworks in Bawtry, Yorkshire. The Environment Agency (EA) served a remediation notice to National Grid Gas Plc, the current owners of the Bawtry site, to pay for cleanup costs. The decision was taken to appeal and overruled on the grounds that National Grid did not "cause or knowingly permit any substances to be in, on or under the land" as stated in Part 2A of the Environmental Protection Act. The damage to the land was actually caused by East Midlands Gas and its predecessors, before National Grid (formerly Transco Plc) was even formed. The appeal was a great victory for National Grid as the ruling could have opened the door to liability cases at numerous gasworks sites owned by the company [R (on the application of National Grid Gas Plc) v. Environment Agency, 2007].

However, not all former gasworks sites are kept as a single property. During the decline of the industry, many sites may have been redeveloped and split into multiple land holdings. In addition, a number of different land uses, including various industrial processes, may have occurred at the sites over the years, leading to multiple sources of contamination. These factors emphasise the importance of accurate chemical fingerprinting to unambiguously prove who is liable for a particular contaminated area.

For environmental forensic evidence to be heard by the court, it must be deemed admissable. As previously described, the USA employs a set of strict rules governing the admissability of scientific evidence, known as the Daubert criteria. However, in the UK there are no set rules, only exclusions based on previous case law [Fraser, 2010]. Ultimately, the decision to reject or accept evidence lies with the judge. It is thought that the lack of strict rules may make it easier to introduce new technology or methods in a courtroom setting. However, a judge who will generally have little or no scientific knowledge must decide if such techniques are reliable and relevant to the case. This can cause a lag period between the introduction of new techniques and the uptake of the technology by the courts [Bertsch, 1999].

The aim of this Chapter is to validate the statistical model for coal tar source apportionment through application of a blind study. Furthermore, the application of ultra resolution chemical fingerprinting as a tool for environmental forensic applications will be evaluated using the Daubert criteria and compared to traditional methods by cost analysis.

### 8.1.1   Aim and Objectives

The overall aim of this chapter was to validate the source allocation model discussed in the previous chapter and to examine the introduction of new analytical techniques (such as GCxGC TOFMS) in a courtroom setting where admissibility of evidence is paramount. To achieve this aim, the following objectives were identified:

- Case study of unknown samples to test the classification power of the coal tar source allocation model.

- The evaluation of the ultra resolution method using a set of criteria generally applied to science in law.

- The cost comparison of ultra resolution chemical fingerprinting with traditional analytical approaches.

## 8.2   Materials and Methods

### 8.2.1   Case Study

Sample preparation, GCxGC TOFMS analyses and statistical data processing of four additional coal tar samples were performed according to the methods described in Chapters 6 and 7 respectively. Samples 20 (split into sub-samples 20t1, 20t3 and 20t5) and 23 were obtained from two additional FMGP sites. The site details were not disclosed by Parsons Brinckerhoff until both analysis and data processing had been completed, to avoid any potential bias in the interpretation of results. Coal tar sample 20 was provided in five separate containers, sampled at the same location. To ensure the method of sampling did not affect the final results, three of the containers were selected

at random and coal tar samples from each were extracted. Due to the high viscosity of certain coal tar samples, it was difficult to thoroughly mix the tar prior to sampling, thus it was difficult to show that an aliquot representative of the entire sample was used. In total, four additional samples were extracted and analysed (in duplicate) and added to the dataset used in Chapter 7 to allow a new PCA model to be created.

## 8.3    Results and Discussion

### 8.3.1    Case Study

The optimal method of statistical processing for GCxGC analyses of coal tar was described in the previous Chapter. A principal component analysis (PCA) score plot was developed with the capacity to classify all major coal tar types investigated in this study. Coal tar samples from two additional FMGP sites, S13 and S15, were analysed for the purpose of a blind study. The historic information for both sites is summarised in Chapter 4, however, this data was not disclosed by Parsons Brinckerhoff until the samples were analysed and evaluated statistically.

The GCxGC chromatograms of the four blind study samples are provided in Figure 8.1. It can be seen that all three samples (20t1, 20t3 and 20t5) from site S13 are very similar and noticeably different from the coal tar from site S15. The samples from S13 appear to have a large range in composition, with many aromatics appearing in high concentration, unlike many of the coal tar samples where only certain PAHs are found in high abundance.
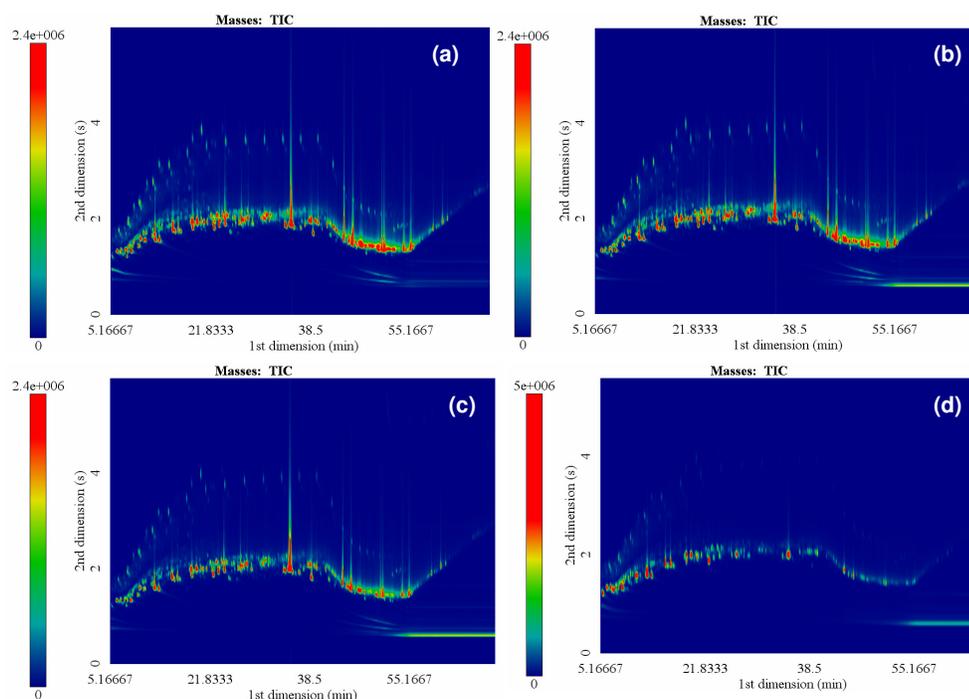
Figure 8.1: Reversed phase GCxGC contour plots of the blind study coal tars (a) 20 t1, (b) 20 t3, (c) 20 t5 and (d) 23.

The optimal univariate statistical method using a double ratio plot was then performed to discover whether the method was capable of correctly classifying the samples; the resulting diagnostic plot is shown in Figure 8.2. The unknown samples from sites S13 and S15 are grouped within the CWG and retort carbonisation categories respectively. However, this is an area with much overlap between classes, therefore a firm classification cannot be made.
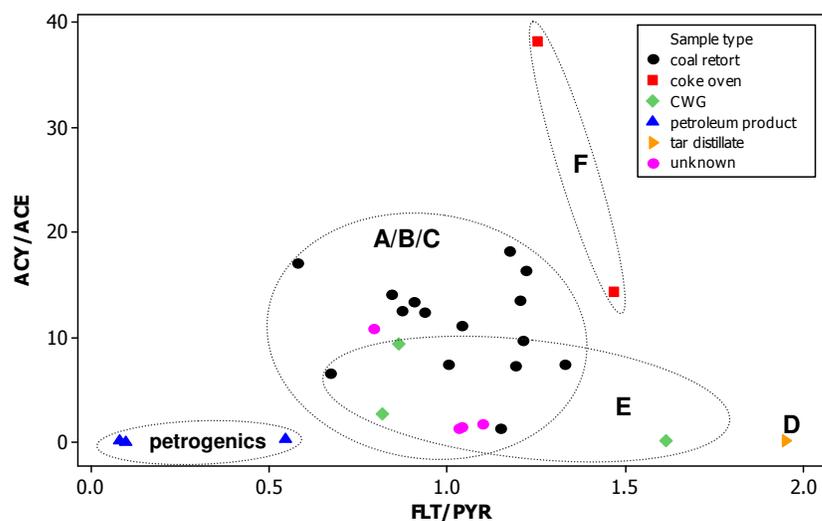
Figure 8.2: Diagnostic ratio plot of the entire coal tar library including the blind study samples. Lettering corresponds to the system adopted in Chapter 7.

In an attempt to provide classification of tars by their manufacturing process, the optimal multivariate method was performed with the inclusion of the additional unknown samples. The resulting PCA score plot is provided in Figure 8.3 with the various manufacturing classes labelled as discussed in Chapter 7.
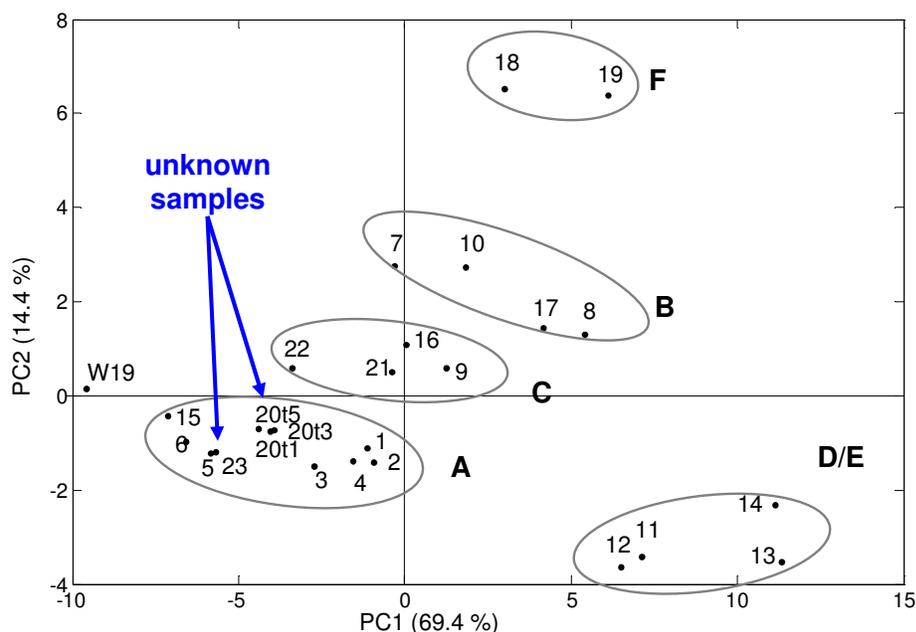
Figure 8.3: PCA score plot of the entire coal tar library including the blind study samples, 20t1, 20t3, 20t5 and 23. A=vertical retort, B=horizontal retort, C=low temperature horizontal retort, D= creosote, E=carburetted water gas and F=coke oven

The score plot in Figure 8.3 shows that samples 20 (t1, t3 and t5) and 23 are all classified as vertical retort tars by PCA. Furthermore, the sub-samples from sites S13 (20t1, 20t3 and 20t5) cluster closely in the score plot, indicating that the sampling method provided a representative portion of the coal tar and did not produce a major effect on the acquired chemical fingerprint. The high viscosity of certain tar samples from the library made it very difficult to homogenise tars within the large (250-1000 mL capacity) containers prior to sampling. It was therefore difficult to ensure that sampling was representative of the entire sample, and it was unknown how this would affect the obtained chemical fingerprint.

The historical site data confirms that sites S13 and S15 did indeed use vertical retort carbonisation. The initial chromatographic pattern differences in samples 20t1, t3 and t5 could be explained by the presence of a naphtha reforming plant at site S13. Naptha reforming is unlikely to have produced tar, however, any naphtha leakages may ultimately dissolve into any tar with which they came into contact. Naphtha is a complex mixture of alkanes, naphthenes and aromatics in the C5-C12 region [Prestvik et al., 2004]. This is consistent with the observed chromatographic patterns, shown in Figures 8.1a, b and c, thus it is possible that these samples represent mixtures of coal tar and

naphtha. The chromatographic differences between these samples and other vertical retort tars (where parent PAHs dominate) may also be due to the presence of additional CWG plants at both sites. However, given the data on the years of operation and production capacity of the retort processes at sites S13 and S15, it is likely that vertical retort tar contamination would be most abundant. For example, site S13 was in production by 1885 using horizontal retort technology, however, in 1920 the process was replaced with vertical retort technology. Between 1920 and 1947 four additional sets of vertical retorts were added to increase gas production levels. Likewise, at site S15, horizontal retorts were originally used from the time of construction of the MGP site, between 1896-1915, until 1916 producing approximately 90.4m cf of gas during this time. Vertical retorts were then installed in 1916 capable of producing 300,000 cf of gas per day and an additional set of vertical retorts added in 1938 to boost gas production by 600,000 cf per day. The production capacity of the vertical retorts used at both sites make it more probable that the tars were sourced from these processes rather than other manufacturing processes present on site.

While chromatographic pattern recognition and univarite statistics failed to provide conclusive classifications of tar type, the PCA model was successful in classifying the unknown samples as being produced by vertical retorts. The case study has therefore validated the model and shown the ease by which samples can be quickly screened by GCxGC and assigned to a particular manufacturing process.

## 8.3.2   Implications of ultra resolution chemical fingerprinting

### Environmental Risk

Remediation of contaminated sites is now a major issue due to the influence of environmental legislation such as the European Liabilities Directive (2004/35/EC). The directive states that polluters must ensure contaminated land is remediated appropriately or provide financial compensation to allow cleanup efforts to be enforced. Robust chemical fingerprinting methods can aid remediation efforts by not only identifying the contaminants which require removal but also with the allocation of blame. Under Part 2A of the Environmental Protection Act in the United Kingdom, liability falls with the person(s) who caused or knowingly permitted a site to become contaminated land. If such a person cannot be identified, liability falls with the current owner of the land. Therefore, it is

highly beneficial for land owners to be able to prove that a prior owner actually contaminated the land. Otherwise they may be left with remediation costs, which can cost many thousands (if not millions) of pounds in the case of large FMGP sites. This illustrates the huge impact which ultra resolution chemical fingerprinting could have for environmental forensic investigations at FMGP sites. The technique has the ability to classify tars based on the particular manufacturing process used to produce them. The contamination must have occurred when this manufacturing process was in use, thus narrowing down the potential parties responsible.

There are a number of previous legal cases involving contamination at FMGPs and similar contaminated sites. In a trial similar to the landmark Corby case in the U.K. (which involved poorly managed remediation efforts at a former steel works) a case was brought against the owners of an FMGP site in Illinois, USA. The Central Illinois Public Service Company (now Ameren CIPS) were alleged to have used inadequate remediation measures while removing 50,000 gallons of buried coal tar [Deardorff, 2002]. The children of four families in the nearby town of Taylorville, Illinois, were diagnosed with rare central nervous system cancers, believed to have been caused by airborne toxins generated during the remediation efforts. In 1998, Ameren CIPS were ordered to pay the four families $3.2 million in compensation. Ameren CIPS appealed the decision, however, the ruling was affirmed by the Illinois Supreme Court [Donaldson v. Central Illinois Public Service Company, 2002]. Given the extreme number of FMGP sites across the globe and the rapid growth of the environmental forensic industry it is reasonable to assume there may be many cases like this in the future. Ultra resolution chemical fingerprinting has the potential to provide comprehensive characterisation of contaminants, thus allowing appropriate risk assessments to be performed to ensure that no toxic chemicals are overlooked [Deardorff, 2002].

## Compliance with Daubert Criteria

For ultra resolution chemical fingerprinting to be used for analysis of commercial samples which could potentially be involved in legal cases, the method must be found admissable in court. In the United Kingdom, there are currently no set rules used to assess the admissability of scientific evidence, the judge will make this decision using a set of exclusions based on previous case law. In the USA, a set of rules known as the Daubert criteria were developed to impose strict regulations on what evidence should be allowed [Daubert v. Merrell Dow Pharmaceuticals, 1993]. The Daubert

criteria is composed of four main questions which are used to assess the admissability of scientific evidence, namely; "(1) Does the theory or technique involve testable hypotheses? (2) Has the theory or technique been subject to peer review and publication? (3) Are there known or potential error rates and are there standards controlling the technique's operation? (4) Is the method or technique generally accepted in the scientific community?" [Brilis et al., 2000].

The ultra resolution chemical fingerprinting technique meets each of these requirements. The technique involves testable hypotheses as the most likely questions involving contaminated land are "where is the source of the contamination?" and "who is reponsible?". Ultra resolution chemical fingerprinting has been shown to be capable of providing the data to answer both of these questions at contaminated sites such as former manufactured gas plants. The method described in this thesis provides fast and effective process-specific allocation of FMGP by-product tars allowing the contamination to be assigned to a particular time period (and thus the responsible operator) when that manufacturing process was employed. Furthermore, the method is capable of differentiating between other anthropogenic sources of PAHs, such as petroleum-derived contamination, which may be present on site due to additional land use since decommissioning of the gasworks as well as from neighbouring sites.

The ultra resolution methodology has been published in a peer-reviewed journal [McGregor et al., 2011a] and the statistical methods used for interpretation of GCxGC data are currently under review in an additional journal paper [McGregor et al., 2011b] thus meets the requirement as set out in the Daubert criteria (see Appendix D for published research). There is a danger with "self-quoting" as an expert witness, however, in the past year two additional research papers have been published (by different authors) on GCxGC analyses of coal tars [Machado et al., 2011, Vasilieva et al., 2011], meaning that the topic is gaining interest and acceptance. In addition, GCxGC technology is well-established across many scientific disciplines, including biological sciences, food technology, forensic science and the petrochemical industry [Marriott and Shellie, 2002, Dalluge et al., 2003], illustrating that it has become accepted by the scientific community. Moreover, numerous manufacturers now produce a range of GCxGC instruments in a competitive market further proving that it is established technology.

The reliability of the method was demonstrated through repeat extraction and analysis of coal tar samples. The results given in Chapter 6 indicate that error rates were acceptable based on US EPA standards. Surrogates were added to monitor the sample preparation process and to ensure that

Table 8.1: Limits of detection of the GCxGC TOFMS instrument for a variety of PAH compunds [adapted from Kerr, 2011].

| Compound | Mass on column (g) |
|---|---|
| naphthalene | $3.33 \times 10^{-10}$ |
| 2-methyl naphthalene | $7.93 \times 10^{-10}$ |
| 1-methyl naphthalene | $6.52 \times 10^{-10}$ |
| 1-ethyl naphthalene | $1.34 \times 10^{-10}$ |
| phenanthrene | $9.67 \times 10^{-11}$ |
| pyrene | $5.57 \times 10^{-11}$ |
| chrysene | $1.31 \times 10^{-10}$ |

extraction efficiency remained within 70-120% the range recommended by the US EPA. Moreover, an internal standard monitored the instrumental precision (including injection efficiency). Ultra resolution chemical fingerprinting involves fewer analytical instruments and less rigorous sample preparation than conventional approaches, thus has a lower potential for both instrumental and human error. The limitations of the procedure (mainly concerning the limit of detection for target compounds) have also been evaluated. Research at the University of Strathclyde performed by Kerr [2011] demonstrated the instrumental limit of detection for a variety of PAH compounds. The results of are summarised in Table 8.1. The detectable mass on column (by GCxGC TOFMS) is consistently within the picogram to femtogram range, adequate for analysis of low concentration components in coal tar extracts.

## Cost Analysis

In commercial environmental forensic laboratories, the cost of an analytical procedure is extremely important, clients may opt for other techniques (or even other labs) if the cost of analysis is too high. As GCxGC TOFMS represents relatively new technology, the instrument price range is very high compared to that of conventional GC techniques. An initial investment of approximately £250,000 must be made for a GCxGC TOFMS, with pricing for GCxGC FID slightly lower due to the lower specificity of the detector. To gauge the approximate cost of GCxGC TOFMS analyses per coal tar sample, cost analysis of the entire method was performed by Victoria Pitstra as part of an MSc project at the University of Strathclyde [Pitstra, 2011]. The cost analysis compared the expense involved in both the ultra resolution method and traditional tiered approaches to chemical fingerprinting. The tiered approach was chosen for comparison as it is currently utilised in oil spill

analysis which targets compounds similar to those in coal tar (with the exception of biomarkers).

It was anticipated that the ultra resolution approach would be far more expensive than analysis using the tiered approach. The additional sample fractionation steps required in the tiered method raise solvent consumption, however, the use of large quantities of liquid nitrogen in GCxGC analysis was expected to offset this saving. Furthermore, the price of the GCxGC and the software required is over £100,000 more than the prices of the GC-FID and GC-MS combined.

To obtain the cost of materials required for extraction of each coal tar, a simple calculation was performed by dividing the unit cost by the number (or amount) of that item required per sample. In order to calculate the cost of the entire method, a rough estimate of the annual sample capacity was calculated. As all three gas chromatographic techniques (GC-FID, GC-MS and GCxGC TOFMS) employed the same temperature programming method the run time for each sample is the same using each of these techniques; approximately 75 minutes (with a faster blank run of 20 minutes between samples). Using the assumption that the instruments can be run 24 hours a day, it is possible to run approximately 15 samples a day (plus blank samples). For approximately 200 working days per year, the total instrumental capacity per year can be calculated at roughly 3000 samples. The estimated lifespan of analytical instruments such as gas chromatographs and mass spectrometers (approximately 10 years) can then be used to calculate the cost per year, including a 10% depreciation value of the instrument (Equation 8.3.1).

$$\text{cost per year} = \frac{\frac{\text{Cost of instrument}}{10 \text{ years}}}{100} \text{ x } 10 \tag{8.3.1}$$

The cost per sample can then be calculated by dividing the cost per year by the total sample capacity per year. The costs of other major pieces of equipment, such as the air compressor, nitrogen generator and autosamplers, associated with the instruments were also calculated in this way. The full economic breakdown of the tiered and ultra resolution approaches are provided in Tables 8.2 and 8.3 respectively. Other running costs such as electricity, estates and overheads for the university were included by addition of 126% per sample. Commercial research performed within any laboratory at the University of Strathclyde is subject to a 35% levy, thus it has also been included in the calculations.

It was discovered that the total cost of using a tiered approach for the analysis of coal tars was approximately £372. However, the conventional method does not always use both the GC-FID

Table 8.2: Breakdown of costs for tiered approach for chemical fingerprinting [taken from Pitstra, 2011].

| Tiered Approach | Materials | Unit Cost | Cost per sample (£) |
|---|---|---|---|
| Extraction and Fractionation | Filter papers | £35 for pack | 0.7 |
| | Diatomaceous Earth | £132.30 per 1kg | 0.13 |
| | NaSO$_4$ | £91.40 per 2.5kg | 0.37 |
| | Silica gel 60 | £56.50 per kg | 1.13 |
| | Hexane | £11.42 per 2.5L | 0.5 |
| | Toluene | £7.24 per 2.5L | 0.17 |
| | Technician Wage | £426.60 per day | 31.00 |
| Recurring Costs | University Levy | 35% of Cost | 11.90 |
| | Electricity/Estates/Overheads | 126% of Cost | 42.84 |
| | **Cost of Extraction** | | **88.74** |
| GC-FID | Instrument | £16,000 | 0.05 |
| | Air | £4.92 per month | 0.01 |
| | Nitrogen | £129.94 per month | 0.29 |
| | Hydrogen | £149.28 per month | 0.33 |
| | Helium | £73 per 3 weeks | 0.23 |
| | Columns | £598.92 per year | 0.2 |
| Interpretation and Reporting | Technician Wage (1 hour per sample) | £426.60 per day | 53.33 |
| Recurring Costs | University Levy | 35% of Cost | 30.94 |
| | Electricity/Estates/Overheads | 126% of Cost | 111.40 |
| | **Cost of GC-FID** | | **142.01** |
| | **Cost of Extraction and GC-FID Analysis** | | **230.75** |
| GC-MS | Helium | £73 per 3 weeks | 0.23 |
| | Columns | £598.92 per year | 0.20 |
| | Instrument | £54,000 | 0.18 |
| | Autosampler | £80,000 | 0.27 |
| Interpretation and Reporting | Technician Wage (1 hour per sample) | £426.60 per day | 53.33 |
| Recurring Costs | University Levy | 35% of Cost | 18.97 |
| | Electricity/Estates/Overheads | 126% of Cost | 68.30 |
| | **Cost of GC-MS** | | **141.48** |
| | **Cost of Extraction and GC-MS Analysis** | | **230.22** |
| | **Total Cost of Extraction, GC-FID Analysis and GC-MS Analysis** | | **372.23** |

and GC-MS. It is important to mention that it is often not necessary to employ both tiers of this approach if the first tier provides conclusive information for source allocation. Generally, both tiers will only be employed if the results are inconclusive or ambiguous, thus significantly affecting the cost. The economic breakdown of the ultra resolution approach, given in Table 8.3, shows the total cost per coal tar sample to be approximately £489, which is £116 more expensive than the conventional method. Commercial analysis of tar samples by a UK provider, using GC-FID and GC-MS, is approximately £170 per sample with an additional fee of around £200 for sample correlation and the generation of a full results report [Thomas, 2011b]. The cost analysis performed for the tiered approach in this project matches well with these figures, however, it may in fact be an overestimate for analysis in a commercial setting, as a higher sample turnover rate (due to fewer sample replicates and blanks) would decrease the overall cost per sample.

The cost of extraction using ultra resolution chemical fingerprinting is approximately £27 cheaper than the traditional methods, providing a 30% saving per sample due to the lack of fractionation for GCxGC. The greatest expenditure for both methods is the salary of trained technical staff to analyse samples and interpret the results. Due to the advanced technology used within the ultra resolution approach, a greater amount of time is required for interpretation of the complex results. The cost of a technician to run and analyse one sample using the tiered approach was calculated at £138 (excluding recurring costs) which can be compared to the cost of £182 for the ultra resolution method. Collation and statistical analysis of GCxGC data takes approximately 3-4 hours and requires a technician trained in the specialised software and statistical methods, while the analysis of data produced by GC-FID and GC-MS takes approximately 1 hour per instrumental technique. This increase in interpretation time cancels out any savings made by faster sample preparation and lower solvent consumption in the ultra resolution approach.

An increase of £116 per sample for the ultra resolution approach would appear to encourage the use of conventional tiered methods, however, the extra charge must be weighed against the added benefits of the technique. Based on the results presented in this thesis, it is reasonable to assume that GC-FID and GC-MS analysis will not provide conclusive classification or source allocation of coal tar samples, thus requiring the use of additional ancillary techniques. The univariate methods employed by Mauro [2000] and Saber et al. [2005] after GC-MS analyses were shown in Chapter 7 to be inadequate for classification of coal tars. The PCA model in this study, however, provides clear, process-specific apportionment of tars using analysis on a single instrument. Furthermore, the £116

additional expenditure is a minute sum compared to the potential compensation claims and legal fees resulting from poor characterisation and remediation of contaminated land. Additionally, the compounds present in coal tars are very similar to those found in crude oils, thus the ultra resolution method may also have potential applications in source provenancing of spilled oils, which is one of the major applications in environmental forensics.

## 8.4   Conclusion

The validation of the principal component model for classification of coal tars was performed through a blind study. Two unknown coal tars were correctly classified after analysis by ASE offline reversed phase GCxGC and subsequent data processing. The classification model illustrates that source-specific and process-specific apportionment of coal tars is possible.

The application of ASE offline reversed phase GCxGC for ultra resolution chemical fingerprinting on a commercial basis was also discussed. The method addresses all the main issues associated with the admissability of scientific evidence in court. The only real disadvantage of the technique, in a courtroom setting, is the increased complexity of the advanced analytical instruments and statistical processing which must be explained in a way non-scientists can comprehend.

Cost analysis of the ultra resolution procedure demonstrated the increased cost compared to analysis using conventional GC instruments in a traditional tiered approach. However, the increased costs must be weighed against the overall benefit of using the ultra resolution method, as it provides greater quantities of chemical information than traditional methods by fast and accurate screening of complex mixtures in a single analytical run.

Table 8.3: Breakdown of costs for ultra resolution chemical fingerprinting [taken from Pitstra, 2011].

| Ultra resolution chemical fingerprinting | Materials | Unit Cost (£) | Cost per sample (£) |
|---|---|---|---|
| Extraction | Filter papers | 35 for pack | 0.7 |
| | Diatomaceous Earth | 132.30 per 1kg | 0.13 |
| | NaSO$_4$ | 91.40 per 2.5kg | 0.37 |
| | Silica gel 60 | 56.50 per kg | 0.17 |
| | 40 mL hexane (ASE) | 11.42 per 2.5L | 0.18 |
| | Technician Wage (25 mins per sample) | £426.60 per day | 22.22 |
| Recurring Costs | University Levy | 35% of Cost | 8.32 |
| | Electricity/Estates/Overheads | 126% of Cost | 29.95 |
| | | **Cost of Extraction** | **62.04** |
| GCxGC TOFMS | Liquid Nitrogen | £77.73 per 35 samples | 2.22 |
| | Helium | £73 per 3 weeks | 0.23 |
| | Columns | £360.59 | 0.12 |
| | GCxGC | £250,000 | 0.83 |
| | Air Compressor | £6,241.21 | 0.02 |
| | Nitrogen Generator | £5,287.00 | 0.02 |
| Interpretation and Reporting | Technician Wage (3hrs per sample) | £426.60 per day | 159.99 |
| Recurring Costs | University Levy | 35% of Cost | 57.20 |
| | Electricity/Estates/Overheads | 126% of Cost | 205.92 |
| | | **Cost of GCxGC TOFMS** | **426.55** |
| | | **Cost of Extraction and GCxGC TOFMS Analysis** | **488.59** |

# Chapter 9

# Conclusions and recommendations

## 9.1 Restatement of the Research Objectives

### 9.1.1 Research Aim and Objectives

The aim of this research is to investigate an efficient form of chemical fingerprinting using advanced analytical techniques to allow more accurate and precise source apportionment and to gain knowledge on the chemical composition of coal tars. To achieve the aims of this project, the following research objectives were identified and met:

- To evaluate the potential of compound specific isotope analysis (CSIA) as a tool for source differentiation of coal tars.

- To optimise the use of two-dimensional gas chromatography (GCxGC) for the analysis of coal tars.

- To develop a fast, accurate and precise extraction procedure to allow analysis of complex coal tars by GCxGC without the need for complex fractionation processes.

- To identify the main classes of chemicals present in coal tars through tentative GCxGC identification for the purpose of chemical fingerprinting.

- To compare statistical methods of processing GCxGC datasets for source apportionment of coal tars.

## 9.2    Summary Details of Key Findings for Chapter 5 to Chapter 8

### 9.2.1    Isotopic Analysis of Coal Tars

A novel automated method for the fractionation of coal tars prior to compound specific isotope analysis (CSIA) was investigated. Accelerated solvent extraction (ASE) was found to produce aromatic coal tar fractions suitable for CSIA, through in-cell cleanup and fractionation by silica gel. Evaluation of the isotopic results, using the Student's T-test, indicated that the majority of isotopic values obtained through ASE fractionation of coal tars were not statistically different to those obtained through traditional silica column chromatography.

The carbon isotopic composition of the aromatic fractions of 25 coal tars was investigated. The majority of coal tar PAHs were found to have carbon isotope values in the narrow range of -24 to -28 ‰, thus it was difficult to assign source based on the isotopic composition. However, the potential of CSIA for FMGP site investigation was evident in multiple samples from across the same site. Samples 1-6 showed much isotopic variation across site S1, with samples 2 and 3 resulting in isotopically light values indicative of petrogenic rather than pyrogenic contamination. This suggests that CSIA could be useful for tracking multiple contaminant plumes across a single FMGP site, but the isotopic signature is not sufficient for full source discrimination of coal tars.

### 9.2.2    Reversed Phase Two-Dimensional Gas Chromatography

Chemical fingerprinting of coal tar by comprehensive two-dimensional gas chromatography (GCxGC) was evaluated. The enhanced resolution of GCxGC allowed the entire coal tar sample to be analysed without the need for time-consuming fractionation processes. An automated extraction method was designed using ASE to provide a single extract containing all chemical classes present in each coal tar sample. Extraction efficiencies ranged between 80-110% and the relative standard deviation (RSD) for the quantification of PAHs across six replicate coal tar extractions was below 12%.

Generally, GCxGC is performed using normal phase; a non-polar primary column connected to a polar secondary column. However, reversing the column configuration has been shown to provide better group-type separation in a few recent literature examples. Optimisation of GCxGC

separation showed an improvement in chromatographic space occupation when using a reversed phase column configuration (81%) compared to the more widely used normal phase mode (76%). Furthermore, the calculated resolution between various PAH isomer pairs was consistently higher for reversed phase compared to normal phase GCxGC.

The repeatability of the reversed phase method was compared to both GC-MS and normal phase GCxGC techniques through quantification of the EPA PAHs present in six replicate injections of a coal tar extract. GC-MS was shown to have poor accuracy due to partially coeluting isomer pairs, with relative standard deviation (RSD) ranging from 0.8 - 17.6%. Normal phase GCxGC provided improved separation, resulting in RSD values below 10% for all investigated PAHs. However, reversed phase GCxGC further improved the repeatability with RSD values below 5%.

The enhanced separation capacity of GCxGC allows a range of chemical classes to be monitored in a single-step analytical procedure, providing an accurate and precise method of ultra resolution chemical fingerprinting for coal tars.

### 9.2.3    Source Apportionment of Coal Tar

The large quantities of chemical data generated by GCxGC analysis of coal tars was compared using a variety of statistical methods. Simple univariate and bivariate statistical methods were shown to be inadequate for the purpose of source allocation, with no decipherable trends in coal tar composition.

The use of multivariate statistical methods, such as hierarchical cluster analysis (HCA) and principal component analysis (PCA), to compare the full coal tar GCxGC datasets of 3495 peak areas was also unsuccessful. The large number of compositional similarities between many of the coal tars appeared to mask out the differences caused by differing manufacturing process. Therefore, a reduced data set containing 156 peak areas was selected and subjected to the same multivariate methods. It was found that preprocessing strategies were essential to normalise the large range in peak intensities across the data set. HCA of the preprocessed, reduced dataset provided improved classification over univariate/bivariate methods, however, similar tars such as those produced by various retort carbonisation processes were often incorrectly classified.

The optimal statistical method for classification of coal tar was found for PCA of the reduced dataset using eighth root preprocessing. The first two principal components explained 83% of the total variation within the data and allowed the classification of five separate coal tar types; namely vertical retorts, horizontal retorts, low temperature horizontal retorts, coke oven and carburetted water gas tar/creosote.

### 9.2.4 Ultra Resolution Chemical Fingerprinting as a Tool for Environmental Forensics

The PCA coal tar classification model, as detailed in Chapter 7, was validated through the case study of two additional FMGP sites. The site details were not disclosed until the analysis and data processing of all coal tar samples was completed. Both sites were correctly classified by the PCA model as having used vertical retort carbonisation processes.

Cost analysis of the ultra resolution chemical fingerprinting method was compared to that of the traditional tiered approaches. It was discovered that the ultra resolution method was £116 more expensive per sample than using traditional approaches. However, the extra expense can be balanced out by the additional information provided by GCxGC. Furthermore, traditional approaches utilise GC-MS and GC-FID which may not be adequate for source identification, thus costly ancillary techniques (such as stable isotope analysis) must also be employed.

## 9.3   Conclusions

Isotopic analysis is thought to be the ideal method for source allocation where contaminant sources are chemically similar. However, the isotopic composition of coal tars in this study were within a narrow range and unable to be apportioned, by neither origin nor specific manufacturing processes. It is believed that for British FMGPs, the extreme similarities between the raw materials (i.e. British coals) used for gas manufacture resulted in coal tars with similar isotopic composition, regardless of differences between the manufacturing processes used in their formation. Isotopic analyses are useful when there is great variety in the type of raw material or, in the case of naturally formed substances, the regional variation (which is affected by isotopic variation in atmospheric

carbon dioxide). This study has demonstrated that ultra resolution GCxGC analysis is a better choice for the source apportionment of complex mixtures when the raw materials and feedstocks are similar. However, compound specific isotope analysis can prove useful in distinguishing differences in contamination across a single site.

This thesis has presented the development of an ultra resolution method of chemical fingerprinting for coal tar samples, which provides new scientific insight into their compositional variation. The ultra resolution method has been shown to provide improved levels of information on coal tar composition than those achieved by conventional approaches. Furthermore, the method benefits from a single-step extraction and analysis on a single analytical instrument using an ASE offline reversed phase GCxGC procedure. This results in fast sample turnover rates and lower risk of both human and instrumental error due to limited sample handling and fewer analytical measurements.

Robust chemical fingerprints of 25 coal tar samples were obtained via the ultra resolution approach. The interpretation of the large quantities of chemical data produced by GCxGC proved extremely challenging without the use of multivariate statistical techniques. Optimisation of statistical interpretation resulted in the development of a classification model capable of process-specific allocation of coal tars, with the potential for site-specific allocation also. This allows the differentiation of multiple coal tar plumes across a single site and the allocation of contamination to a manufacturing process and the specific time period, and thus operator(s) responsible for the tar contamination. Furthermore, the model has the potential to discriminate between coal tars and contamination which is unrelated to FMGP sites, thus preventing the incorrect allocation of blame for instances of pollution.

## 9.4   Recommendations for future work

The reduced set of 156 peaks selected from the GCxGC chromatograms for statistical analyses provide process-specific apportionment of FMGP by-product tars. However, this results in a large portion of the chromatographic data being discarded. Any major differences in chromatographic pattern can be identified visually, however, it is not known if any important trace level variables are being overlooked. Therefore, it would be beneficial to develop a system for the automated comparison of datasets as new samples are added to the model, to highlight major variations in composition

and ensure these variables are also incorporated into the statistical model. This is currently limited by the GCxGC software capabilities. Furthermore, the method of manually selecting the peak areas for statistical comparison is time-consuming and raises costs for possible commercial applications of the method. The development of a data processing procedure to automatically search and compile the peak area data for the 156 selected compounds would be extremely useful as it would save time, and thus reduce the costs associated with interpretation of results.

The statistical model produced for the coal tar library has been shown to provide process-specific classification of coal tars, however, it is not known how this method would cope with extremely large datasets. For example, if the coal tar library was expanded to include worldwide tar samples, it may be necessary to research an alternative method. An interesting possibility would be the use of artificial neural network (ANN) software, such as Kohonen self-organising feature maps (SOFM). The benefit of SOFM using artificial neural networks is the training mechanism used to "teach" the network to recognise specific patterns in a dataset. Subsequent datasets will then be classified according to the pre-existing trained model. The use of SOFM was investigated for this study, though it was found that the coal tar library was too small for the technique to be useful. The technique has been shown to be useful for the differentiation of ignitable liquids in arson investigation [Mat Desa et al., 2010], however, SOFM uses a form of unsupervised learning, so the more information used to train the model the greater the ability of the technique for classification purposes. Consequently, an expanded version of the coal tar library would greatly benefit from SOFM data processing.

The expansion of the coal tar library to include worldwide samples could provide further information on compositional changes due to differences in both the manufacturing processes and raw materials used in their production. This could allow greater accuracy in determining the operator(s) responsible for tar contamination at an FMGP site. Furthermore, it would be interesting to include a greater number of samples from other common types of contaminated land to test the capability of the statistical model to differentiate between the sources. This could potentially provide a single statistical model with the capacity to differentiate between many types of contamination encountered in environmental forensic investigations. The classification model could be expanded to include common petrogenic and pyrogenic PAH contamination sources to allow background contamination or multiple contamination sources at/near FMGP sites to be apportioned and to ensure the blame is not solely allocated to former gas-making processes.

The PCA score plot used to approximate weathering in Chapter 7 is another area of research which could be expanded upon. Only one fresh tar sample was weathered to aid interpretation of the plot. It would be interesting to include chromatographic data from fresh coal tars which have been subjected to varying degrees of air evaporation. Additionally, experiments to imitate subsurface water-washing could be performed on fresh tars to provide another layer of information on the compositional changes caused by weathering. In turn, this may aid the source apportionment studies, by ensuring that coal tars are not incorrectly classified due to extreme alteration of chemical fingerprints by weathering.

To the author's knowledge, there are currently no reports in the literature of the hydrogen isotope composition of coal tars. Hydrogen isotope analysis is a less robust method than that of carbon CSIA, however, the combination of carbon and hydrogen isotope data may aid the source differentiation process, especially as hydrogen isotope values are known to show greater variation than those of carbon. It is possible that 2D isotopic data (i.e. carbon and hydrogen) could provide verification of petrogenic mixing within coal tar plumes at FMGP sites.

# References

Daubert v. Merrell Dow Pharmaceuticals, 509, U.S. 579, 1993.

Donaldson v. Central Illinois Public Service Company, Illinois Supreme Court, 199 Ill. 2d 63, 89, 2002.

Corby Group Litigation v Corby District Council, EWHC 2109, 2009.

Frye v. United States, 293 F. 1013, 1014, D.C. cir. 1923.

T. A. T. Aboul Kassim and B. R. T. Simoneit. Petroleum Hydrocarbon Fingerprinting and Sediment Transport Assessed by Molecular Biomarker and Multivariate Statistical Analyses in the Eastern Harbour of Alexandria, Egypt. *Marine Pollution Bulletin*, 30:63–73, 1995.

M. Adahchour, J. Beens, R. J. J. Vreuls, and U. A. Brinkman. Recent developments in comprehensive two-dimensional gas chromatography (GCxGC): I. Introduction and instrumental set-up. *Trends in Analytical Chemistry*, 25(5):438–454, 2006a.

M. Adahchour, J. Beens, R. J. J. Vreuls, and U. A. Brinkman. Recent developments in comprehensive two-dimensional gas chromatography (GCxGC): II. Modulation and detection. *Trends in Analytical Chemistry*, 25(6):540–553, 2006b.

M. Adahchour, J. Beens, R. J. J. Vreuls, and U. A. Brinkman. Recent developments in comprehensive two-dimensional gas chromatography (GCxGC): III. Applications for petrochemicals and organohalogens. *Trends in Analytical Chemistry*, 25(7):726–741, 2006c.

H. Alimi, T. Ertel, and B. Schug. Fingerprinting of Hydrocarbon Fuel Contaminants: Literature Review. *Environmental Forensics*, 4:25–38, 2003.

ASTM. D3328, Standard Test Methods for Comparison of Waterborne Petroleum Oils by Gas Chromatography. Technical report, American Society for Testing and Materials, 1974.

K. Benhabib, P. Faure, M. Sardin, and M. Simonnot. Characteristics of a solid coal tar sampled from a contaminated soil and of the organics transferred into water. *Fuel*, 89:352–359, 2010.

S. Benson, C. Lennard, P. Maynard, and C. Roux. Forensic applications of isotope ratio mass spectrometry - A review. *Forensic Science International*, 157:1–22, 2006.

W Bertsch. Two-Dimensional Gas Chromatography. Concepts, Instrumentation, and Applications- Part 1: Fundamentals, Conventional Two-dimensional Gas Chromatography, Selected Applications. *Journal of High Resolution Chromatography*, 22(12):647–665, 1999.

P.S. Birak and C.T. Miller. Dense non-aqueous phase liquids at former manufactured gas plants : Challenges to modeling and remediation. *Journal of Contaminant Hydrology*, 105:81–98, 2009.

P. D. Boehm, G. S. Douglas, W. A. Burns, P. J. Mankiewicz, D. S. Page, and A. E. Bence. Application of Petroleum Hydrocarbon Chemical Fingerprinting and Allocation Techniques after the Exxon Valdez Oil Spill. *Marine Pollution Bulletin*, 34(8):599–613, 1997.

P. D. Boehm, D. S. Page, E. S. Gilfillan, A. E. Bence, and P. J. Burns, W. A.and Mankiewicz. Study of the Fates and Effects of the Exxon Valdez Oil Spill on Benthic Sediments in Two Bays in Prince William Sound, Alaska. 1. Study Design, Chemistry, and Source Fingerprinting. *Environmental Science and Technology*, 32:567–576, 1998.

C.R. Bozek. Yesterday's gas, today's coal tar. *Electric Perspectives*, 20:38–42, 1995.

J. T. Brenna, T. N. Corso, H. J. Tobias, and R. J. Caimi. High-precision continuous flow isotope ratio mass spectrometry. *Mass Spectrometry Reviews*, 16:227–258, 1997.

G. M. Brilis, J. C. Worthington, and A. D. Wait. Quality Science in the Courtroom: U.S. EPA Data Quality and Peer Review Policies and Procedures Compared to the Daubert Factors. *Environmental Forensics*, 1:197–203, 2000.

D.G. Brown, L. Gupta, T.H. Kim, H.K. Moo Young, and A.J. Coleman. Comparative assessment of coal tars obtained from 10 former manufactured gas plant sites in the Eastern United States. *Chemosphere*, 65:1562–1569, 2006.

229

P. Burchill, A. A. Herod, and E. Pritchard. Investigation of nitrogen-containing compounds in coal tar products. 1. Unfractionated materials. *Fuel*, 62:11–19, 1983.

J. W. A. Butterfield. *Chemistry of Gas Manufacture, Volume I Materials and Processes.* Charles Griffin and Company, London, 3rd edition, 1904.

J.F. Carter, R. Sleeman, J.C. Hill, F. Idoine, and E.L. Titterton. Isotope ratio mass spectrometry as a tool for forensic investigation (examples from recent studies). *Science and Justice*, 45(3): 141–149, 2005.

M. D. Casal, M. A. Diez, R. Alverez, and C. Barriocanal. Primary tar of different coking coal ranks. *International Journal of Coal Geology*, 76:237–242, 2008.

M. R. Cave, J. Wragg, I Harrison, C. H. Vane, T. Van De Wiele, E. De Grove, C. P. Nathanail, M. Ashmore, R. Thomas, J. Robinson, and P. Daly. Comparison of Batch Mode and Dynamic Physiologically Based Bioaccessibility Tests for PAHs in Soil Samples. *Environmental Science and Technology*, 44(7):2654–2660, 2010.

CERCLA. Comprehensive Environmental Response, Compensation and Liability Act. P.L. 96-510, 42 U.S.C., December 1980.

J. H. Christensen, A. B. Hansen, G. Tomasi, J. Mortensen, and O. Andersen. Integrated methodology for forensic oil spill identification. *Environmental Science and Technology*, 38:2912–2918, 2004.

J. H. Christensen, G. Tomasi, and A. B Hansen. Chemical Fingerprinting of Petroleum Biomarkers Using Time Warping and PCA. *Environmental Science and Technology*, 39:255–260, 2005.

J.H. Christensen and G. Tomasi. Practicl aspects of chemometrics for oil spill fingerprinting. *Journal of Chromatography A*, 1169:1–22, 2007.

F. D'Affonseca, P. Blum, M. Finkel, R. Melzer, and P. Grathwohl. Field scale characterization and modeling of contaminant release from a coal tar source zone. *Journal of Contaminant Hydrology*, 102:120–139, 2008.

P.S. Daling, L.G. Faksness, A.B. Hansen, and S.A. Stout. Improved and standardized methodology for oil spill fingerprinting. *Environmental Forensics*, 3:263–278, 2002.

J. Dalluge, J. Beens, and U.A. Brinkman. Comprehensive two-dimensional gas chromatography: a powerful and versatile analytical tool. *Journal of Chromatography A*, 1000:69–108, 2003.

E. de Hoffman and V. Stroobant. *Mass Spectrometry: Principles and Applications*. John Wiley & Sons, 3rd edition, 2007.

S. de Koning. GCxGC(-ToF MS) Fundamentals and Advanced Method Development, Leco European GCxGC workshop, Amsterdam. 2009.

J. Deardorff. Cancer victims' long fight ends. Technical report, Chicago Tribune, 2nd April, http://articles.chicagotribune.com [as accessed on 05/10/2011], 2002.

Department of Communities and Local Government (DCLG). Securing the Future Supply of Brownfield Land, Government Response to English Partnerships' Recommendations on the National Brownfield Strategy, (http://www.communities.gov.uk/documents/planningandbuilding/doc/securingfuturebrownfield.doc as accessed on 20/09/11). Technical report, 2008.

Department of Environment, Food and Rural Affairs,. Directive 2004/35/CE of the European Parliament and Council of the European Union on environmental liability with regard to the prevention and remedying of environmental damage, March 2009.

Department of the Environment. Gas Works, Coke Works and Other Coal Carbonisation Plants. Technical report, Industry Profile, UK., 1995.

G. S. Douglas, S. Emsbo Mattingly, S. A. Stout, A. D. Uhler, and K. J. McCarthy. Chemical Fingerprinting Methods. In B.L. Murphy and R.D Morrison, editors, *Introduction to Environmental Forensics*, pages 312–439. Elsevier, London, 2nd edition, 2007.

S. Emsbo Mattingly and P. Boehm. Identifying PAHs from Manufactured Gas Plant Sites. Technical Report 1005289, Electric Power Research Institute (EPRI), 2003.

S. D. Emsbo Mattingly, A. D. Uhler, S. A Stout, G. S. Douglas, K. J. McCarthy, and A. Coleman. Determining the source of PAHs in sediment. *Land Contamination and Reclamation*, 14(2): 403–411, 2006.

Environmental Protection Act. Part 2A, Contaminated Land Regime, 1990.

C. S. Eskilsson and E. Bjorklund. Analytical-scale microwave-assisted extraction. *Journal of Chromatography A*, 902:227–250, 2000.

L. S. Ettre. Nomenclature for chromatography (IUPAC Recommendations 1993). *Pure and Applied Chemistry*, 65(4):819–972, 1993.

B. Everitt and T. Hothorn. *An Introduction to Applied Multivariate Analysis with R*. Springer, 2011.

B. S. Everitt and G. S. Dunn. *Multivariate Data Analysis*. Hodder Arnold, 2001.

R. Fernandez Varela, J. M. Andrade, S. Muniategui, and D. Prada. Selecting a reduced suite of diagnostic ratios calculated between petroleum biomarkers and polycylic aromatic hydrocarbons to characterize a set of crude oils. *Journal of Chromatography A*, 1217:8279–8289, 2010.

A. Findlay. *The Treasures of Coal Tar*. Turnbull and Spears, U.K., 1917.

W. Francis and M.C. Peters, editors. *Fuels and Fuel Technology*. Pergamon Press, Oxford, second edition, 1980.

J. Fraser. *Forensic Science: A Very Short Introduction*. Oxford University Press, U.K., 2010.

G.S. Frysinger and R.B. Gaines. Forensic analysis of ignitable liquids in fire debris by comprehensive two-dimensional gas chromatography. *Journal of Forensic Sciences*, 47(3):471–482, 2002.

G.S. Frysinger, R.B. Gaines, and C.M. Reddy. GCxGC - A New Analytical Tool For Environmental Forensics. *Environmental Forensics*, 3:27–34, 2002.

R.B. Gaines, G.S. Frysinger, M.A. Hendrick Smith, and J. D. Stuart. Oil spill source identification by comprehensive two-dimensional gas chromatography. *Environmental Science and Technology*, 33(12):2106–2112, 1999.

R.B. Gaines, G.S. Frysinger, C.M. Reddy, and R.K. Nelson. Oil spill source identification by comprehensive two-dimensional gas chromatography (GCxGC). In Z. Wang and S. A. Stout, editors, *Oil Spill Environmental Forensics: Fingerprinting and Source Identification*, pages 169–206. Elsevier, London, 2007.

W. M. Gardner, editor. *The British Coal-Tar Industry*. William and Norgate, London, 1915.

C. Gauchotte, G. O'Sullivan, S. Davis, and R.M. Kalin. Development of an advanced on-line position-specific stable carbon isotope system and application to methyl tert-butyl ether. *Rapid Communications in Mass Spectrometry*, 23:3183–3193, 2009.

C. Gauchotte, A. Assal, L. A. McGregor, and R. M. Kalin. One-step Accelerated Solvent Extraction of Soil and Coal Tar Samples for Compound Specific Isotope Analysis of PAHs. In *International Network of Environmental Forensics (INEF) conference, Cambridge, UK, poster presentation*, 2011.

O. Gibert, A.S. Ferguson, R.M. Kalin, R. Doherty, K.W. Dickson, K.L. McGeough, J. Robinson, and R. Thomas. Performance of a Sequential Reactive Barrier for Bioremediation of Coal Tar Contaminated Groundwater. *Environmental Science and Technology*, 41:6795–6801, 2007.

M. C. Graham, R. Allan, A. E. Fallick, and J. G. Farmer. Investigation of extraction and clean-up procedures used in the quantification and stable isotopic characterisation of PAHs in contaminated urban soils. *Science of the Total Environment*, 360:81–89, 2006.

J. Gray, G. Lacrampe Couloume, D. Gandhi, K. M. Scow, R. D. Wilson, D. M. Mackay, and B. Sherwood Loller. Carbon and Hydrogen Isotopic Fractionation during Biodegradation of Methyl tert-Butyl Ether. *Environmental Science and Technology*, 36:1931–1938, 2002.

R. L. Grob and E. F. Barry, editors. *Modern Practice of Gas Chromatography*. John Wiley and Sons, USA, 4th edition, 2004.

J.A. Hall, R.M. Kalin, M.J. Larkin, C.C.R. Allen, and D.B. Harper. Variation in stable carbon isotope fractionation during aerobic degradation of phenol and benzoate by contaminant degrading bacteria. *Organic Geochemistry*, 30:801–811, 1999.

B. T. Hammer, C. A. Kelley, R. B. Coffin, L. A. Cifuentes, and J. G. Mueller. d13C values of polycyclic aromatic hydrocarbons collected from two creosote-contaminated sites. *Chemical Geology*, 152:43–58, 1998.

M. J. Hamper. Manufactured Gas History and Processes. *Environmental Forensics*, 7:55–64, 2006.

S. M. Harkins, R. S. Truesdale, R. Hill, P. Hoffman, and S. Winters. U.S. Production of Manufactured Gases:Assessment of Past Disposal Practices. Technical report, Research Triangle Institute, North Carolina, prepared for the U.S. Environmental Protection Agency, 1988.

J. Harynuk, T. Gorecki, and J. de Zeeuw. Overloading the second-dimension column in comprehensive two-dimensional gas chromatography. *Journal of Chromatography A*, 1071:21–27, 2005.

A. W. Hatheway. Long-term fate and transport characteristics of manufactured ga plant residuals and wastes. In *International Association for Engineering Geology (IAEG), paper number 566*, 2006.

A.W. Hatheway. Geoenvironmental protocol for site and waste characterization of former manufactured gas plants; worldwide remediation challenge in semi-volatile organic wastes. *Engineering Geology*, 64:317–338, 2002.

A.W. Hatheway and B.C. Doyle. Technical history of the town gas plants of the British Isles, 2006.

J. Hoefs. *Stable Isotope Geochemistry*. Springer, 2nd edition, 1980.

D. Hunkeler, R. U. Meckenstock, B. Sherwood Lollar, T. C. Schmidt, and J.T. Wilson. A Guide for Assessing Biodegradation and Source Identification of Organic Ground Water Contaminants Using Compound Specific Isotope Analysis (CSIA). Technical report, US Environemental Protection Agency, 2008.

J. V. Ibarra, I. Cervero, and R. Moliner. Structural Characterization of Low Temperature Pyrolysis Tars and their Relation with Parent Coals. *Fuel Processing Technology*, 22:135–149, 1989.

ITOPF. Oil Tanker Spill Statistics. Technical report, International Tanker Owners Pollution Federation Ltd. URL http://www.itopf.com/information-services/data-and-statistics/statistics/documents/Statspack2009-FINAL.pdf [as accessed on 15/09/2011], 2009.

B. Jones and N. Parpworth. *Environmental Liabilities*. Shaw, Crayford, 2004.

F. Keppler, R. M. Kalin, D. B. Harper, W. C. McRoberts, and J. T. G. Hamilton. Carbon isotope anomaly in the major plant C1 pool and its global biogeochemical implications. *Biogeosciences*, 1:123–131, 2004.

S. Kerr. Monitoring of chemical oxidation remediation of polycyclic aromatic hydrocarbons by comprehensive two-dimensional gas chromatography. Master's thesis, MRes, University of Strathclyde, Glasgow, 2011.

C. A. King. *King's Manual of Gas Manufacture*. Walter King Ltd., London, 1948.

J. Krupcik, P. Majek, R. Gorovenko, P. Sandra, and D. W. Armstrong. On retentivity tuning by flow in the second column of different comprehensive two dimensional gas chromatographic configurations. *Journal of Chromatography A*, 1218:3186–3189, 2011.

D. W. Later, M. L. Lee, K. D. Bartle, R. C. Kong, and D. L. Vassilaros. Chemical Class Separation and Characterization of Organic Compounds in Synthetic Fuels. *Analytical Chemistry*, 53:1612–1620, 1981.

J. J. Leary and R. L. Schmidt. Quadrupole mass spectrometers: an intuitive look at the math. *Journal of Chemical Education*, 73:1142–1144, 1996.

Leco. *Pegasus HT and ChromaTOF Software Instruction Manual, version 4.0x*. Leco Corporation, 2008.

Leco. Pegasus 4D GCxGC TOFMS brochure, Form No. 209-183. Technical report, 2011.

V. B. Lewes. *The Carbonisation of Coal*. John Allan and Company, London, 1912.

M. E. Machado, E. B. Caramao, and C. A. Zini. Investigation of sulphur compounds in coal tar using monodimensional and comprehensive two-dimensional gas chromatography. *Journal of Chromatography A*, 1218:3200–3207, 2011.

G. Malatesta. *Coal tars and their derivatives*. E. & F. N. Spon, London, 1920.

P. Marriott and R. Shellie. Principles and applications of comprehensive two-dimensional gas chromatography. *Trends in Analytical Chemistry*, 21(9-10):573–583, 2002.

W.N. S. Mat Desa, N. NicDaeid, I. Dzulkiflee, and K. Savage. Application of unsupervised chemometric analysis and self-organising feature map (SOFM) for the classification of lighter fuels. *Analytical Chemistry*, 82:6395–6400, 2010.

D. Mauro. Chemical Source Attribution at Former MGP Sites. Technical Report 1000728, Electric Power Research Institute (EPRI), 2000.

L. Mazeas and H. Budzinski. Molecular and Stable Carbon Isotopic Source Identification of Oil Residues and Oiled Bird Feathers Sampled along the Atlantic Coast of France after the Erika Oil Spill. *Environmental Science and Technology*, 36:130–137, 2002.

L. A. McGregor, C. Gauchotte Lindsay, N. Nic Daéid, R. Thomas, P. Daly, and R. M. Kalin. Ultra resolution chemical fingerprinting of dense non-aqueous phase liquids from manufactured gas plants by reversed phase comprehensive two-dimensional gas chromatography. *Journal of Chromatography A*, 1218:4755–4763, 2011a.

L. A. McGregor, C. Gauchotte Lindsay, N. Nic Daéid, R. Thomas, and R. M. Kalin. Multivariate statistical methods for the environmental forensic classification of coal tars from former manufactured gas plants. *submitted to Environmental Science and Technology*, 2011b.

H. M. McNair and J. M. Miller. *Basic Gas Chromatography*. Wiley, New Jersey, 2nd edition, 2009.

C. McRae, G.D. Love, I.P. Murray, C.E. Snape, and A.E. Fallick. Potential of gas chromatography isotope ratio mass spectrometry to source polycyclic aromatic hydrocarbon emissions. *Analytical Communications*, 33:331–333, 1996.

C. McRae, C. Sun, C. E. Snape, A. E. Fallick, and D. Taylor. d13C values of coal-derived PAHs from different processes and their application to source apportionment. *Organic Geochemistry*, 30:881–889, 1999.

A. Meade. *New Modern Gasworks Practice*, volume 1. Eyre and Spottiswoode, London, 1934.

W. Meier Augenstein. Applied gas chromatography coupled to isotope ratio mass spectrometry. *Journal of Chromatography A*, 842:351–371, 1999.

C. A. Menzie, B. B. Potocki, and J. Santodonato. Exposure to carcinogenic PAHs in the environment. *Environmental Science and Technology*, 26(7):1278–1284, 1992.

G. S. Metcalf. An investigation of oxygen-containing compounds - phenols and indanols - in coal liquid distillates. MSc thesis. Master's thesis, University of Utah, 1983.

B.S. Mitrevski, P. Wilairat, and P.J. Marriott. Comprehensive two-dimensional gas chromatography improves separation and identification of anabolic agents in doping control. *Journal of Chromatography A*, 1217(1):127–135, 2010.

L. Mondello, P. Quinto Tranchida, P. Dugo, and G. Dugo. Comprehensive Two-Dimensional Gas Chromatography-Mass Spectrometry: A Review. *Mass Spectrometry Reviews*, 27:101–124, 2008.

R.D. Morrison. Critical Review of Environmental Forensic Techniques: Part I. *Environmental Forensics*, 1:157–173, 2000.

S. M. Mudge. Environmental Forensics. In R. E. Hester and R. M. Harrison, editors, *Issues in Environmental Science and Technology*, volume 26, pages 1–16. Royal Society of Chemistry, London, 2008.

C. Muhlen, C. A. Zini, E. B. Caramao, and P. J. Marriot. Applications of comprehensive two-dimensional gas chromatography to the characterization of petrochemical and related samples. *Journal of Chromatography A*, 1105:39–50, 2006.

B. L. Murphy, T. Sparacio, and W. J. Shields. Manufactured Gas Plants - Processes, Historical Development, and Key Issues in Insurance Coverage Disputes. *Environmental Forensics*, 6: 161–173, 2005.

N. Nic Daeid, H. A. S. Buchanan, K. A. Savage, J. G. Fraser, and S. L. Cresswell. Recent Advances in the Application of Stable Isotope Ratio Analysis in Forensic Chemistry. *Australian Journal of Chemistry*, 63(1):3–7, 2010.

M. Novotny, J. W. Strand, S. L. Smith, D. Wiesler, and F. J. Schwende. Compositional Studies of Coal Tar by Capillary Gas Chromatography/Mass Spectrometry. *Fuel*, 60:213–220, 1981.

B. Omais, M. Courtiade, N. Charon, D. Thiebaut, A. Quignard, and M. Hennion. Investigating comprehensive two-dimensional gas chromatography conditions to optimize the separation of oxygenated compounds in a direct coal liquefaction middle distillate. *Journal of Chromatography A*, 1218:3233–3240, 2011.

V.P. O'Malley, T.A. Abrajano Jr., and J. Hellou. Determination of the 13C/12C ratios of individual PAH from environmental samples: can PAH sources be apportioned? *Organic Geochemistry*, 21 (6/7):809–822, 1994.

R. Ong, S. Lundstedt, P. Haglund, and P. Marriott. Pressurised liquid extraction-comprehensive two-dimensional gas chromatography for fast-screening of polycyclic aromatic hydrocarbons in soil. 1019(1-2):221–232, 2003.

G. O'Sullivan and R.M. Kalin. Investigation of the range of carbon and hydrogen isotopes within a global set of gasolines. *Environmental Forensics*, 9:166–176, 2008.

O. Panic and T. Gorecki. Comprehensive two-dimensional gas chromatography (GCxGC) in environmental analysis and monitoring. *Analytical and Bioanalytical Chemistry*, 386:1013–1023, 2006.

J. B. Phillips and E. B. Ledford. Thermal modulation: A chemical instrumentation component of potential value in improving portability. *Field Analytical Chemistry and Technology*, 1:23–29, 1996.

J.B. Phillips and J. Beens. Comprehensive two-dimensional gas chromatography: a hyphenated method with strong coupling between the two dimensions. *Journal of Chromatography A*, 856 (1-2):331–347, 1999.

R. P. Philp. The emergence of stable isotopes in environmental and forensic geochemistry studies: a review. *Environmental Chemistry Letters*, 5:57–66, 2007.

R.P. Philp, J.D. Allen, and T. Kuder. The use of the isotopic composition of individual compounds for correlating spilled oils and refined products in the environment with suspected sources. *Environmental Forensics*, 3:341–348, 2002.

V. Pitstra. An investigation into the uptake of new technologies in Environmental Forensics with a case study of a new chemical fingerprinting technique. Master's thesis, MSc, University of Strathclyde, Glasgow, 2011.

R. Prestvik, K. Moljord, K. Grande, and A. Holman. *Catalytic Naphtha Reforming*, chapter Naphtha Reforming Chemistry, pages 1–34. CRC Press, USA, 2nd edition edition, 2004.

D. R. Pryde. *Gas world, coking sections*, 101:108–110, 1934.

Restek. Restek capillary column installation guide. Technical report, Restek Corp, Bellefont,PA, USA, 2008.

E. O. Rhodes. *National Research Council Committee. Chemistry of Coal Utilization*, volume II, chapter The Chemical Nature of Coal Tar, pages 1287–1370. John Wiley and Sons, 1945.

P. Richards, C. Gauchotte Lindsay, R. M. Kalin, L. A. McGregor, H. Keenan, R. Thomas, and P. Daly. A One-Step Method for Priority Compounds of Concern in Tar from Contaminated Land Sites: Derivatisation with GCxGC TOFMS analysis. *submitted to Analytical Chemistry*, 2011.

B. E. Richter, B. A. Jones, J. L. Ezzell, N. L. Porter, N. Avdalovic, and C. Pohl. Accelerated Solvent Extraction: A Technique for Sample Preparation. *Analytical Chemistry*, 68:1033–1039, 1996.

T. Romanowski, W. Funcke, I. Grossmann, J. Koenig, and E. Balfanz. Gas chromatographic mass spectrometric determination of high-molecular-weight polycyclic aromatic hydrocarbons in coal tar. *Analytical Chemistry*, 55:1030–1033, 1983.

D. Saber, D. Mauro, and T. Sirivedhin. Environmental Forensics Investigation in Sediments near a Former Manufactured Gas Plant. *Environmental Forensics*, 7:65–75, 2006.

D.L. Saber, D. Mauro, and T. Sirivedhin. Applications of forensic chemistry to environmental work. *Journal of Industrial Microbiology and Biotechnology*, 32:665–668, 2005.

T. C. Schmidt, L. Zwank, M. Elsner, M. Berg, R. U. Meckenstock, and S. B. Haderlein. Compound-specific stable isotope analysis of organic contaminants in natural environments: a critical review of the state of the art, prospects, and future challenges. *Analytical and Bioanalytical Chemistry*, 378:283–300, 2004.

R. Scott. *Principles and Practice of Chromatography*. www.library4science.com, 2003.

J. V. Seeley, F. Kramp, and C. J. Hicks. Comprehensive Two-Dimensional Gas Chromatography via Differential Flow Modulation. *Analytical Chemistry*, 72:4346–4352, 2000.

A.L. Sessions. Isotope-ratio detection for gas chromatography. *Journal of Separation Science*, 29: 1946–1961, 2006.

E. Skoczyska, P. Korytar, and J. De Boer. Maximizing chromatographic information from environmental extracts by GCxGC-ToF-MS. *Environmental Science and Technology*, 42:6611–6618, 2008.

G. F. Slater. Stable Isotope Forensics - When Isotopes Work. *Environmental Forensics*, 4:13–23, 2003.

B.J. Smallwood, R.P. Philp, and J.D. Allen. Stable carbon isotopic composition of gasolines determined by isotope ratio monitoring gas chromatography mass spectrometry. *Organic Geochemistry*, 33:149–159, 2002.

R. P. Soule. *Examination of Low-Temperature Coal Tar*. PhD thesis, Columbia University, New York, 1922.

S. A. Stout, B. Liu, G. C. Millner, D. Hamlin, and E. Healey. Use of Chemical Fingerprinting to Establish the Presence of Spilled Crude Oil in a Residential Area Following Hurricane Katrina, St. Bernard Parish, Louisiana. *Environmental Science and Technology*, 41:7242–7251, 2007.

S.A. Stout, A.D. Uhler, K.J. McCarthy, and S. Emsbo-Mattingly, editors. *Chemical fingerprinting of hydrocarbons.* Introduction to Environmental Forensics. Elsevier, 2nd edition, 2007.

J.A. Tarr. *Manufactured Gas, History of,*, volume 4 of *Encyclopedia of Energy.* Academic Press, 2004.

J. Terrace. *Terrace's Notebook for Gas Engineers and Students.* Ernest Benn Ltd, London, 1948.

ThermoScientific. Conflo IV Operating Manual. Technical report, 2007.

ThermoScientific. GC-Isolink User Manual. Thermo Scientific Corporation, Massachussets, USA. Technical report, 2008.

ThermoScientific. European Training Institute (ETI), Thermo Fisher Scientific, DSQ Operations Manual. Technical report, 2009.

A. O. Thomas and J. N. Lester. The reclamation of disused gasworks sites: new solutions to an old problem. *The Science of the Total Environment*, 152:239–260, 1994.

R. Thomas. Forensic Investigation of Coal Tar. Technical Report FSE97140C, Parsons Brinckerhoff, June 2011a.

R. Thomas. Parsons Brinckerhoff, personal communication (email), 3rd November. 2011b.

P. Thorsheim. The paradox of smokeless fuels: gas, coke, and the environment in Britain, 1813-1949. *Environment and History*, 8:381–401, 2002.

B. A. Tomkins, H. Kubota, W. H. Griest, J. E. Caton, B. R. Clark, and M. R. Guerin. Determination of Benzo[a]pyrene in Petroleum Substitutes. *Analytical Chemistry*, 52:1331–1334, 1980.

T. C. Tran, G. A. Logan, E. Grosjean, J. Harynuk, D. Ryan, and P. J. Marriot. Comparison of column phase configurations for comprehensive two dimensional gas chromatographic analysis of crude oil and bitumen. *Organic Geochemistry*, 37:1190–1194, 2006.

T. C. Tran, G. A. Logan, E. Grosjean, D. Ryan, and P. J. Marriot. Use of comprehensive two-dimensional gas chromatography/time-of-flight mass spectrometry for the characterization of biodegradation and unresolved complex mixtures in petroleum. *Geochimica et Cosmochimica Acta*, 74:6468–6484, 2010.

R. M. Uhler, E. M. Healey, K. J. McCarthy, A. D. Uhler, and S. A. Stout. Molecular Fingerprinting of Gasoline by a Modified EPA 8260 Gas Chromatography-Mass Spectrometry Method. *International Journal of Environmental Analytical Chemistry*, 83(3):1–20, 2002.

USEPA. Test Methods for Evaluating Solid Wastes, SW-846, Chapter 1: Quality Control, 1992. URL `http://www.epa.gov/waste/hazard/testmethods/sw846/pdfs/chap1.pdf`.

USEPA. Test Methods for Evaluating Solid Wastes, SW-846 Method 3580, Waste Dilution, 1997a.

USEPA. Test Methods for Evaluating Solid Wastes, SW-846 Method 8000B, Determinative Chromatographic Separations, 1997b.

USEPA. Test Methods for Evaluating Solid Wastes, SW-846, Method 3545A, Pressurised Fluid Extraction (PFE), 2007.

USEPA. SW-846 Online Test Methods for Evaluating Solid Waste, Physical/Chemical Methods. Technical report, http://www.epa.gov/waste/hazard/testmethods/sw846/index.htm as accessed on 20th October, 2011a.

USEPA. Priority Pollutants. http://water.epa.gov/scitech/methods/cwa/pollutants.cfm as accessed on 12th july. Technical report, 2011b. URL `http://water.epa.gov/scitech/methods/cwa/pollutants.cfm`.

R. Van Der Westhuizen, A. Crouch, and P. Sandra. The use of GCXGC with time-of-flight mass spectrometry to investigate dienes and Diels Alder polymerisation products in high-temperature Fischer-Tropsch-based fuels. *Journal of Separation Science*, 31:3423–3428, 2008.

V. G. Van Mispelaar, A. K. Smilde, O. E. De Noord, J. Blomberg, and P. J. Schoenmakers. Classification of highly similar crude oils using data sets from comprehensive two-dimensional gas chromatography and multivariate techniques. *Journal of Chromatography A*, 1096:156–164, 2005.

V. Vasilieva, K. E. Scherr, E. Edelmann, M. Hasinger, and A. P. Loibner. Comprehensive GC2/MS for the monitoring of aromatic tar oil constituents during biodegradation in a historically contaminated soil. *Journal of Biotechnology*, Article In Press, 2011.

C. A. G. Veras, J. A. Carvalho, and M. A. Ferreira. The Chemical Percolation Devolatilization Model Applied to the Devolatilization of Coal in High Intensity Acoustic Fields. *Journal of the Brazilian Chemical Society*, 13:358–367, 2002.

Z. Wang and M. Fingas. Developments in the analysis of petroleum hydrocarbons in oils, petroleum products and oil-spill-related environmental samples by gas chromatography. *Journal of Chromatography A*, 774:51–78, 1997.

Z. Wang and S. A. Stout. Chemical fingerprinting of spilled or discharged petroleum - methods and factors affecting petroleum fingerprints in the environment. In Z. Wang and S. A. Stout, editors, *Oil Spill Environmental Forensics: Fingerprinting and Source Identification*, pages 1–53. Elsevier, London, 2007.

Z. Wang and S. A. Stout. Environmental Forensics. In R. E. Hester and R. M. Harrison, editors, *Issues in Environmental Science and Technology*, volume 26, page 54. Royal Society of Chemistry, London, 2008.

Z. Wang, M. Fingas, and D.S. Page. Oil spill identification. *Journal of Chromatography A*, 843: 369–411, 1999.

Z. Wang, M. Fingas, and L. Sigouin. Using multiple criteria for fingerprinting unknown oil samples having very similar chemical composition. *Environmental Forensics*, 3:251–262, 2002.

Z. Wang, M. Fingas, P. Lambert, G. Zengb, C. Yanga, and B. Hollebone. Characterization and identification of the Detroit River mystery oil spill (2002). *Journal of Chromatography A*, 1038: 201–214, 2004.

Z. Wang, C. Yang, M. Fingas, B. Hollebone, U. Yim, and J. Oh. Petroleum biomarker fingerprinting for oil spill characterization and source identification. In Z. Wang and S. A. Stout, editors, *Oil Spill Environmental Forensics: Fingerprinting and Source Identification*, pages 73–146. Elsevier, London, 2007.

A. R. Warne. *Coal Tar Distillation and working up of tar products*. John Allan and Company, London, 1907.

M. Wehrer, T. Rennert, T. Mansfeldt, and K. U. Totsche. Contaminants at Former Manufactured Gas Plants: Sources, Properties, and Processes. *Critical Reviews in Environmental Science and Technology*, 41:1883–1969, 2011.

L. Wilkinson and M. Friendly. The History of the Cluster Heat Map. *The American Statistician*, 63:179–184, 2009.

R. Williams. *A Practical Treatise on the Manufacture and Distribution of Coal Gas*. E. & F. N. Spon, London, 1877.

C. Wittman. Fluxome analysis using GC-MS. *Microbial Cell Factories*, 6(6), 2007.

S. Young. *Distillation Principles and Processes*. MacMillan and Company, London, 1922.

# Appendix A

# Appendices for Chapter 4

## A.1  Dionex ASE 350
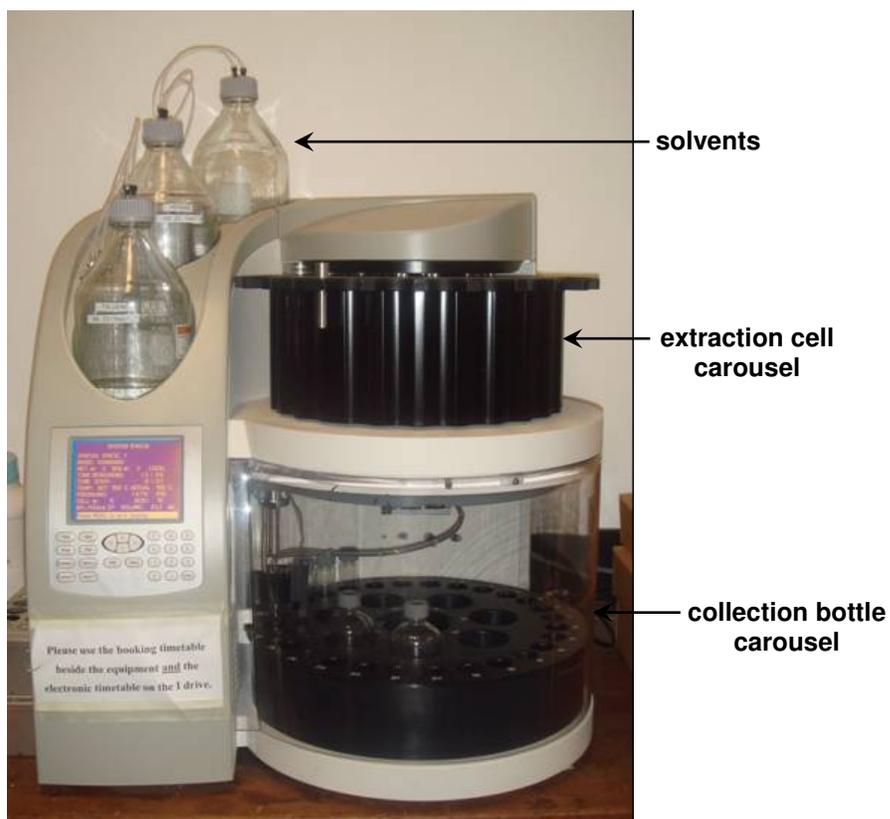


Figure A.1: Dionex ASE 350 accelerated solvent extraction system

## A.2 Büchi Syncore® Analyst
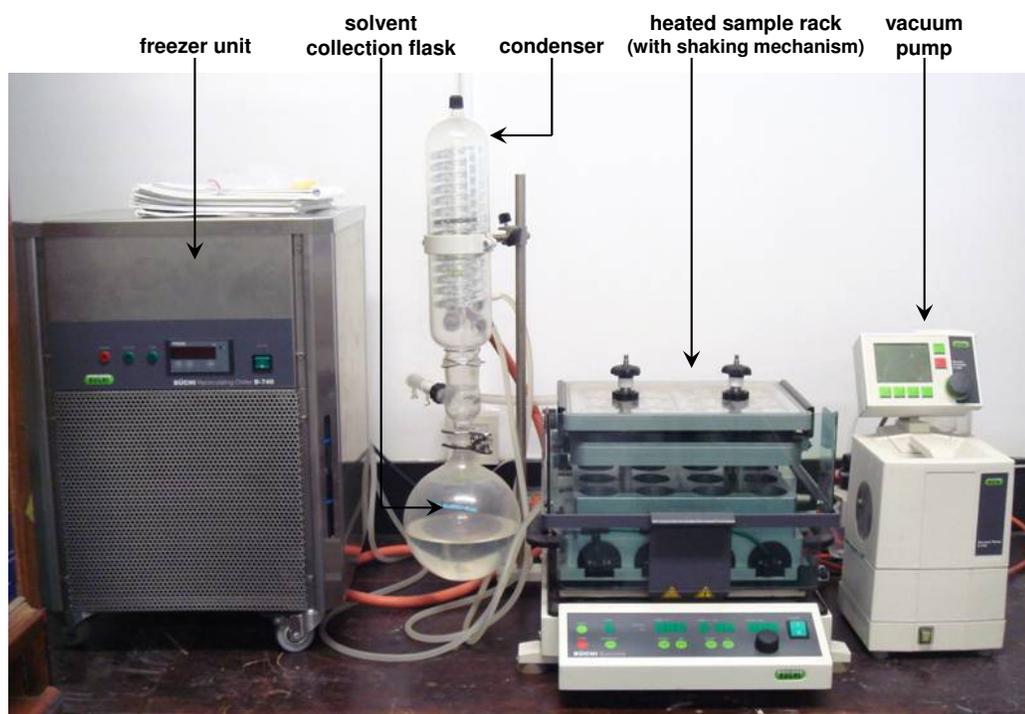


Figure A.2: Büchi Syncore® Analyst

## A.3 Thermo Scientific Trace GC-MS
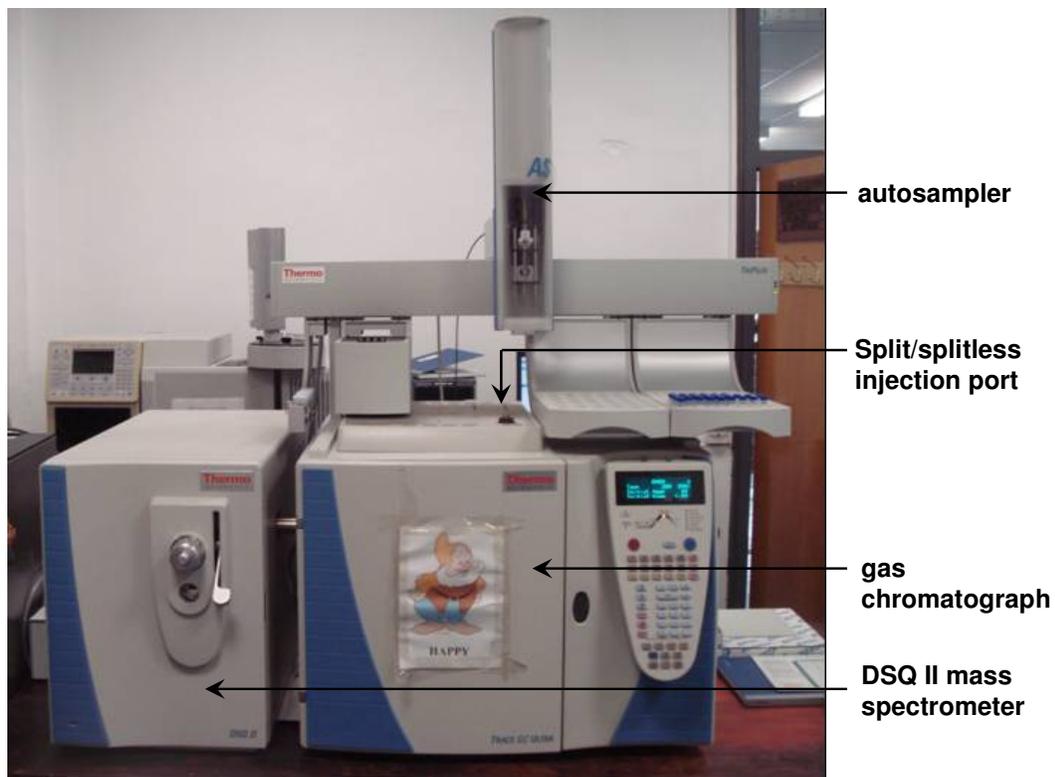


Figure A.3: Thermo Scientific GC - DSQII quadrupole MS

# A.4 Leco Pegasus 4D GCxGC TOFMS
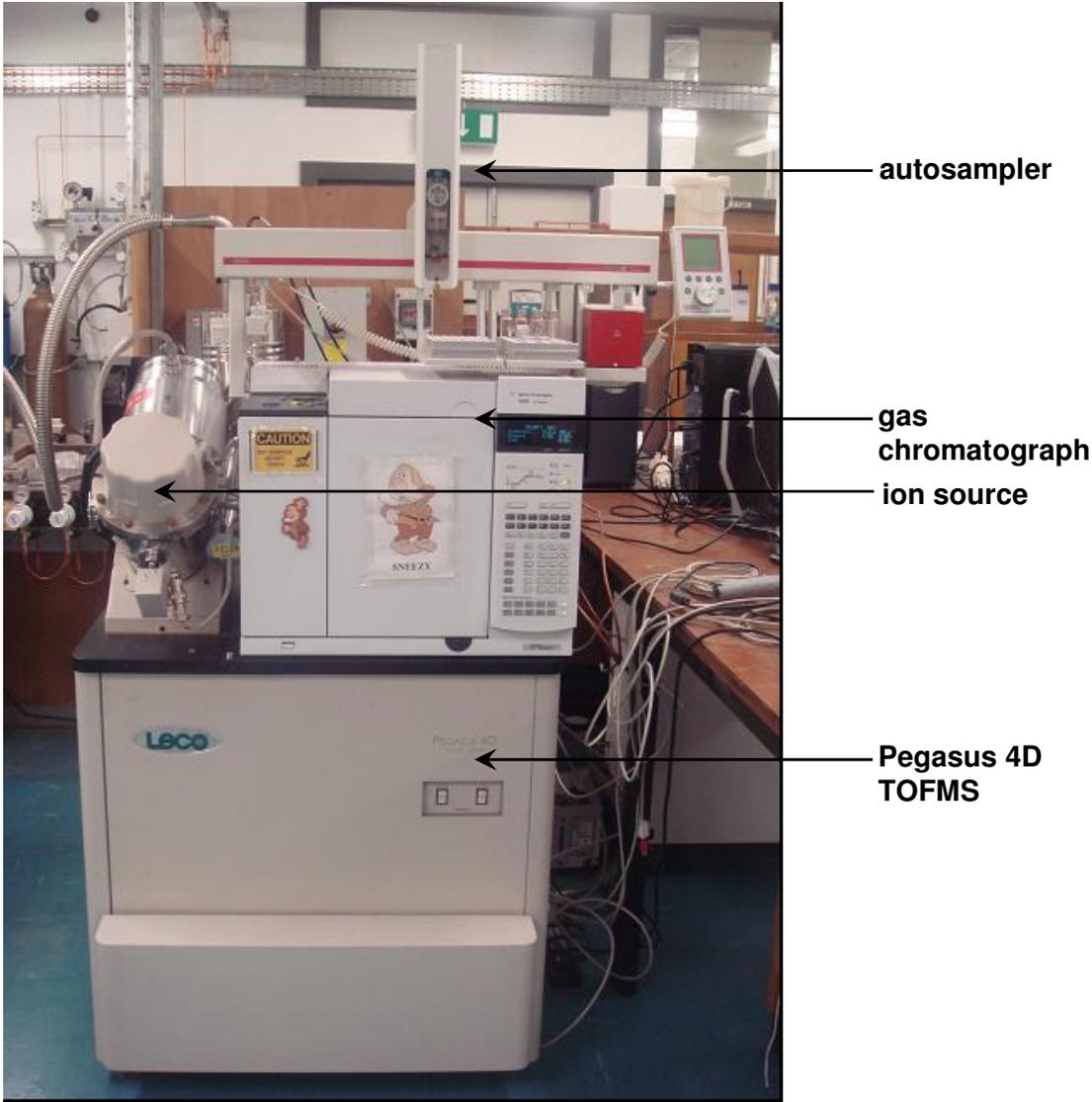


Figure A.4: Full GCxGC TOFMS instrument
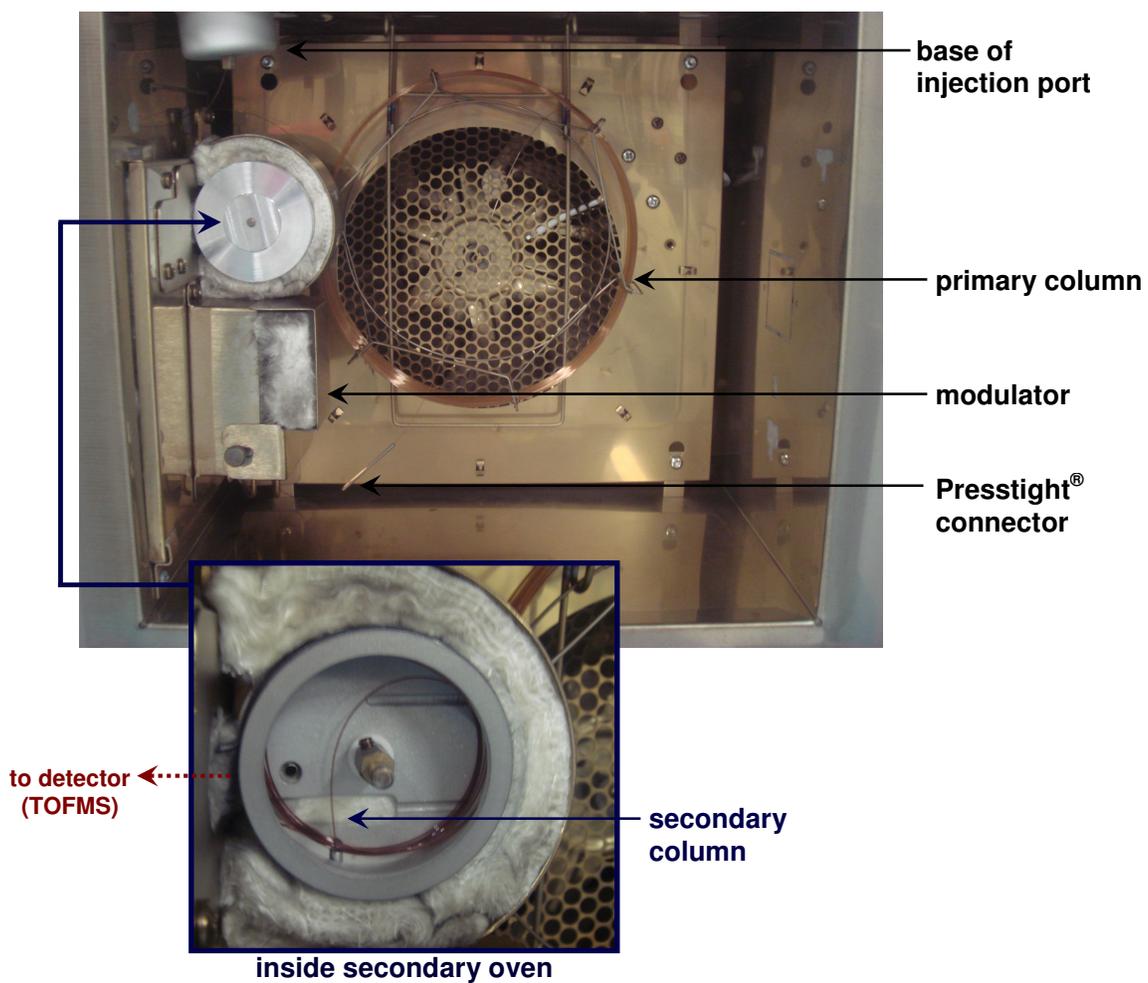
Figure A.5: Inside GCxGC primary oven, including close-up of secondary oven

Figure A.6: Close up of GCxGC cryogenic modulator



Figure A.7: Close-up view of transfer line in the GCxGC TOFMS instrument

## A.5   Thermo Scientific GC Isolink - Delta IV IRMS



Figure A.8: Thermo Scientific GC-C-IRMS

## A.6   Example set of PAH calibration curves

**d8-naphthalene**

140311_PAH_400_2

140311_PAH_250_1

140311_PAH_100_1

140311_PAH_50_1

140311_PAH_25_2

140311_PAH_2.5_1

y= -7.22145x² + 61492.1x - 50606.8
r = 0.99954

**naphthalene**

140311_PAH_400_2

140311_PAH_250_1

140311_PAH_100_1

140311_PAH_50_1

140311_PAH_25_1

140311_PAH_2.5_1

y= -70.0033x² + 126727x - 147579
r = 0.99977

**acenaphthylene**

140311_PAH_400_2

140311_PAH_250_1

140311_PAH_100_1

140311_PAH_50_1

140311_PAH_2.5_1

y= +44.4554x² + 39421.1x - 108769
r = 0.99742

252

**acenaphthene**

140311_PAH_400_2

140311_PAH_250_1

140311_PAH_100_1

140311_PAH_50_1

140311_PAH_2.5_1

y= +7.80851x² + 31997.5x - 11248.7
r = 0.99975



**d10-fluorene**

140311_PAH_400_2

140311_PAH_400_3

140311_PAH_250_1

140311_PAH_100_1

140311_PAH_50_2

140311_PAH_25_1

140311_PAH_2.5_1

y= +2.89409x² + 8257.37x - 7631.3
r = 0.99833



**fluorene**

140311_PAH_400_2

140311_PAH_250_1

140311_PAH_100_1

140311_PAH_50_1

140311_PAH_25_1

140311_PAH_2.5_2

y= +8.83047x² + 33068x - 49936.3
r = 0.99970

phenanthrene

140311_PAH_400_2

140311_PAH_250_1

140311_PAH_100_1

140311_PAH_50_1

140311_PAH_25_1

140311_PAH_2.5_1

$y= +8.65987x^2 + 35877.9x - 13630.5$
$r = 0.99911$



anthracene

140311_PAH_400_2
140311_PAH_400_3

140311_PAH_250_2
140311_PAH_250_1

140311_PAH_100_1

140311_PAH_25_2

140311_PAH_2.5_1

$y= +13.942x^2 + 29025.4x + 12502.7$
$r = 0.99881$



d10-fluoranthene

140311_PAH_400_2

140311_PAH_250_1

140311_PAH_100_1

140311_PAH_50_1

140311_PAH_25_1

140311_PAH_2.5_1

$y= -4.04692x^2 + 14623.5x + 5389.29$
$r = 0.99954$

fluoranthene

140311_PAH_400_2

140311_PAH_250_1

140311_PAH_100_1

140311_PAH_50_1

140311_PAH_25_1

140311_PAH_2.5_1

Area

Concentration

$y = -6.54611x^2 + 29351.3x + 42081.2$
$r = 0.99961$

pyrene

140311_PAH_400_2

140311_PAH_250_1
140311_PAH_250_2

140311_PAH_100_1

140311_PAH_50_1

140311_PAH_25_1

140311_PAH_2.5_1

Area

Concentration

$y = +26195.9x + 29322.1$
$r = 0.99868$

Benz[a]anthracene

140311_PAH_400_3
140311_PAH_400_2

140311_PAH_250_1

140311_PAH_100_1

140311_PAH_50_1

140311_PAH_2.5_1

Area

Concentration

$y = +16.5844x^2 + 9128.71x + 14240.1$
$r = 0.99936$

**d12-chrysene**

140311_PAH_400_2

140311_PAH_250_1

140311_PAH_100_1

140311_PAH_50_1

140311_PAH_25_1

140311_PAH_2.5_1

Area

Concentration

$y = -0.501136x^2 + 7563.73x + 9929$
$r = 0.99932$



**chrysene**

140311_PAH_400_2

140311_PAH_250_1

140311_PAH_100_1

140311_PAH_50_1

140311_PAH_2.5_1

Area

Concentration

$y = +6.38797x^2 + 13907.5x - 4394.83$
$r = 0.99783$



**Benzo[b]fluoranthene**

140311_PAH_400_2
140311_PAH_400_3

140311_PAH_250_1

140311_PAH_100_1

140311_PAH_50_1

140311_PAH_2.5_1

Area

Concentration

$y = +17.694x^2 + 14061.4x - 2685.92$
$r = 0.99797$

256

Benzo[k]fluoranthene

y= +8.95751x² + 7717.93x + 13703.1
r = 0.99901

Benzo[a]pyrene

y= +34.5003x² + 2343.19x - 389.425
r = 0.99951

Indeno[1,2,3-cd]pyrene

y= +23.3509x² + 7658.43x + 6841.44
r = 0.99851

257

Figure A.9: Calibration curves for the quantification of 16 EPA PAHs (and surrogates) by GCxGC TOFMS.

# Appendix B

# Appendices for Chapter 5

Table B.1: Comparison of certified $\delta^{13}C$ values with experimental $\delta^{13}C$ values for replicate analysis of an isotopic reference material (after reference gas correction).

| Alkane | Certified $\delta^{13}C$ value (‰) | Experimental $\delta^{13}C$ values (‰) | | | | | | Mean | Standard Deviation |
|---|---|---|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 | 6 | | |
| C16 | -30.66 | -29.93 | -30.27 | -30.15 | -30.72 | -30.37 | -30.52 | -30.32 | 0.28 |
| C17 | -31.16 | -31.02 | -30.84 | -30.89 | -30.95 | -30.90 | -30.81 | -30.90 | 0.08 |
| C18 | -31.11 | -30.76 | -30.75 | -30.82 | -30.86 | -30.96 | -31.22 | -30.89 | 0.18 |
| C19 | -33.17 | -33.08 | -32.96 | -33.07 | -32.94 | -33.21 | -33.21 | -33.08 | 0.12 |
| C20 | -32.35 | -32.36 | -32.31 | -32.45 | -32.21 | -32.68 | -32.56 | -32.43 | 0.17 |
| C21 | -29.10 | -29.30 | -28.55 | -28.55 | -28.19 | -28.44 | -28.69 | -28.62 | 0.37 |
| C22 | -32.87 | -32.02 | -32.42 | -32.80 | -32.93 | -32.69 | -32.24 | -32.52 | 0.35 |
| C23 | -31.77 | -31.38 | -31.71 | -31.70 | -31.44 | -31.41 | -31.82 | -31.58 | 0.19 |
| C24 | -33.34 | -33.34 | -33.35 | -33.49 | -33.20 | -33.47 | -33.54 | -33.40 | 0.13 |
| C25 | -28.48 | -29.01 | -28.63 | -28.64 | -28.41 | -28.87 | -28.81 | -28.73 | 0.21 |
| C26 | -33.03 | -32.40 | -31.88 | -32.73 | -31.93 | -32.25 | -32.60 | -32.30 | 0.35 |
| C27 | -29.56 | -30.05 | -30.13 | -29.61 | -29.25 | -29.67 | -30.00 | -29.78 | 0.34 |
| C28 | -32.21 | -32.61 | -32.35 | -32.52 | -31.88 | -32.38 | -32.61 | -32.39 | 0.28 |
| C29 | -30.07 | -30.69 | -30.15 | -30.51 | -30.11 | -30.60 | -30.51 | -30.43 | 0.24 |
| C30 | -29.86 | -30.57 | -30.52 | -30.81 | -30.10 | -30.42 | -30.41 | -30.47 | 0.23 |

Table B.2: Mean $\delta^{13}$C values for analysis of in-house PAH standards prior to each sample sequence. Mean and standard deviation of isotopic values for each PAH across the three month period are also provided.

| Date | F8-N | N | Acy | Ace | Flu | Phe | Ant | Flt | Pyr | BaA | Chr | BbF+BkF | BaP | IP + DBA | BP |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 4/1/2011 | -21.8 | -26.1 | -24.6 | -24.3 | -27.3 | -24.9 | -26.2 | -25.1 | -26.0 | -26.9 | -26.3 | -26.4 | -31.0 | -26.1 | -29.8 |
| 4/2/2011 | -21.8 | -26.7 | -24.9 | -24.6 | -27.6 | -25.3 | -25.7 | -25.3 | -25.8 | -26.8 | -26.4 | -26.7 | -31.1 | -26.0 | -29.8 |
| 4/3/2011 | -21.2 | -26.0 | -24.7 | -24.9 | -27.2 | -24.5 | -25.6 | -24.8 | -26.5 | -26.1 | -26.3 | -26.5 | -30.6 | -26.9 | -29.7 |
| 4/5/2011 | -21.0 | -26.6 | -25.0 | -24.3 | -27.7 | -25.7 | -26.2 | -26.2 | -26.6 | -27.1 | -26.6 | -27.0 | -31.8 | -26.9 | -30.8 |
| 4/6/2011 | -20.6 | -26.5 | -25.3 | -25.2 | -27.7 | -25.2 | -26.3 | -25.7 | -25.6 | -26.5 | -26.7 | -27.0 | -31.3 | -26.1 | -30.2 |
| 4/7/2011 | -21.6 | -26.1 | -24.7 | -24.5 | -27.4 | -25.0 | -25.5 | -25.2 | -25.2 | -26.4 | -26.6 | -26.8 | -31.1 | -26.3 | -29.6 |
| 4/10/2011 | -21.4 | -26.5 | -24.9 | -24.9 | -27.6 | -25.4 | -25.8 | -26.0 | -26.1 | -26.6 | -27.1 | -27.0 | -32.1 | -27.4 | -31.6 |
| 4/14/2011 | -22.3 | -26.7 | -24.8 | -25.0 | -27.8 | -25.0 | -26.5 | -25.6 | -25.5 | -25.9 | -27.0 | -27.0 | -31.5 | -26.7 | -30.1 |
| 4/19/2011 | -21.7 | -26.8 | -24.6 | -24.4 | -27.2 | -24.8 | -26.2 | -26.1 | -26.0 | -26.4 | -26.6 | -26.4 | -31.6 | -26.8 | -29.8 |
| 4/22/2011 | -22.1 | -27.0 | -25.5 | -25.0 | -27.7 | -25.9 | -26.7 | -26.1 | -26.0 | -26.7 | -27.6 | -27.8 | -31.8 | -26.4 | -30.2 |
| 5/20/11 | -22.3 | -26.6 | -24.7 | -24.4 | -27.2 | -24.7 | -25.6 | -25.3 | -25.2 | -26.5 | -26.6 | -26.7 | -31.7 | -25.7 | -29.9 |
| 5/23/11 | -22.2 | -27.2 | -25.8 | -25.7 | -28.2 | -25.7 | -26.7 | -26.0 | -26.7 | -27.1 | -26.8 | -27.4 | -31.8 | -26.6 | -29.8 |
| 5/28/11 | -22.4 | -27.4 | -25.6 | -25.3 | -27.8 | -25.7 | -26.9 | -26.3 | -27.0 | -27.0 | -27.3 | -27.6 | -32.1 | -27.3 | -30.6 |
| 5/29/11 | -22.4 | -27.5 | -25.4 | -25.1 | -28.1 | -25.8 | -25.9 | -25.9 | -26.1 | -27.7 | -26.5 | -28.0 | -32.7 | -27.1 | -29.5 |
| 6/4/2011 | -22.2 | -26.9 | -25.2 | -24.7 | -27.8 | -25.1 | -27.1 | -25.8 | -26.3 | -26.7 | -27.7 | -27.3 | -31.5 | -26.9 | -29.9 |
| 6/6/2011 | -22.5 | -26.9 | -25.2 | -25.1 | -28.3 | -25.5 | -26.3 | -25.9 | -26.6 | -26.4 | -27.7 | -27.8 | -32.2 | -27.0 | -30.5 |
| mean | -21.8 | -26.7 | -25.1 | -24.8 | -27.7 | -25.3 | -26.2 | -25.7 | -26.1 | -26.7 | -26.9 | -27.1 | -31.6 | -26.6 | -30.1 |
| std dev | 0.55 | 0.44 | 0.39 | 0.38 | 0.34 | 0.42 | 0.48 | 0.46 | 0.53 | 0.42 | 0.49 | 0.52 | 0.52 | 0.49 | 0.53 |

Table B.3: Carbon isotopic values of PAHs in coal tar (from fraction 2 of each sample).

| Sample | Mean $\delta^{13}$C values (‰)[a] | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | N | 2-MeN | 1-MeN | Acy | Ace | Flu | Phe | Ant | Flt | Pyr | BaA | Chr | BbF + BkF | BaP | IP + DBA | BP |
| 1 | -26.5 | -25.9 | -26.2 | -25.7 | -26.2 | -26.9 | -27.2 | -27.1 | -26.4 | -26.8 | -25.2 | -26.1 | -24.8 | -25.7 | low[b] | low[b] |
| 2 | -31.0 | -29.5 | -27.7 | -30.2 | -29.5 | -28.8 | -30.2 | -27.6 | -29.9 | -30.1 | -29.3 | -29.1 | -29.3 | -29.6 | -28.4 | -28.9 |
| 3 | -30.3 | -32.1 | -30.7 | -26.3 | -27.3 | -29.9 | -29.2 | -27.5 | -27.7 | -28.4 | -28.1 | -28.0 | -27.8 | -27.6 | -26.5 | -26.4 |
| 4 | -26.9 | -28.1 | -27.3 | -24.2 | -26.6 | -26.3 | -26.5 | -27.4 | -27.1 | -28.0 | -26.4 | -27.3 | -27.8 | -26.4 | -28.7 | -29.0 |
| 5 | -26.7 | -26.7 | -26.3 | -26.7 | -25.5 | -26.5 | -26.5 | -27.3 | -26.7 | -27.4 | -26.9 | -27.1 | -26.9 | -27.2 | -26.3 | -26.1 |
| 6 | -26.6 | -28.3 | -27.1 | -25.7 | -25.7 | -26.4 | -27.0 | -25.1 | -26.1 | -26.8 | -25.7 | -26.5 | -26.0 | -25.5 | -27.3 | -25.3 |
| 7 | -25.0 | -25.4 | -24.5 | -23.3 | -24.6 | -25.2 | -26.0 | -24.9 | -26.1 | -25.8 | -25.3 | -25.1 | -25.4 | -25.4 | -24.9 | -23.4 |
| 8 | -25.6 | -25.7 | -26.7 | -24.6 | -25.3 | -26.4 | -26.0 | -26.4 | -26.2 | -26.6 | -25.7 | -26.0 | -25.7 | -25.6 | -25.5 | -26.2 |
| 9 | -25.6 | -26.8 | -26.4 | -23.4 | -26.5 | -26.2 | -26.5 | -25.9 | -26.2 | -26.3 | -25.8 | -25.7 | -26.2 | -24.3 | -25.2 | -26.7 |
| 10 | -26.3 | -26.8 | -25.9 | -26.3 | -27.3 | -28.7 | -27.4 | -27.5 | -28.1 | -27.7 | -29.2 | -27.5 | -28.7 | -28.1 | -29.1 | -28.8 |
| 11 | -27.4 | -22.7 | -24.7 | low[b] | -25.5 | -24.4 | -25.4 | -25.7 | -25.1 | -25.5 | -23.8 | -25.0 | -25.5 | -24.7 | -24.2 | -26.8 |
| 12 | -28.2 | -26.1 | -25.0 | -24.4 | -24.7 | -25.6 | -25.9 | -25.8 | -26.1 | -25.8 | -25.1 | -26.1 | -25.3 | -25.4 | -26.1 | -24.7 |
| 13 | -25.9 | -25.7 | -26.2 | -23.7 | -23.7 | -26.8 | -27.1 | -27.1 | -27.6 | -27.8 | -26.7 | -27.6 | -27.3 | -27.8 | -26.1 | -26.1 |
| 14 | -25.8 | -26.3 | -25.8 | -23.7 | -24.0 | -27.0 | -26.8 | -25.6 | -26.4 | -26.5 | -25.3 | -26.2 | -25.6 | -25.9 | -25.9 | -25.8 |
| 15 | -27.2 | low[b] | low[b] | -28.4 | -27.2 | -29.0 | -28.9 | -27.2 | -29.0 | -29.4 | -30.4 | -28.4 | -29.1 | -29.9 | -29.4 | -29.0 |
| 16 | -25.7 | -25.0 | -25.3 | -23.6 | -24.2 | -25.6 | -26.3 | -25.6 | -26.2 | -26.2 | -26.3 | -25.9 | -26.2 | -26.0 | -25.7 | -26.5 |
| 17 | -25.6 | -25.5 | -25.8 | -23.8 | -24.0 | -26.2 | -26.2 | -25.9 | -26.4 | -26.5 | -25.4 | -25.7 | -25.8 | -25.7 | -25.8 | -26.5 |
| 18 | -26.1 | -25.0 | -24.0 | -25.4 | -24.7 | -26.6 | -26.6 | -26.1 | -26.8 | -26.7 | -27.9 | -26.6 | -26.7 | -26.7 | -26.3 | -27.2 |
| 19 | -26.5 | -25.7 | -25.6 | -23.5 | -25.5 | -27.4 | -28.0 | -24.6 | -27.3 | -27.0 | -27.5 | -26.6 | -26.9 | -27.1 | -26.6 | -26.6 |
| 20t1 | -26.9 | -27.2 | -24.9 | -25.5 | -27.6 | -28.0 | -27.6 | -25.8 | -27.7 | -28.0 | -27.0 | -25.6 | -26.4 | -26.1 | -25.4 | -27.0 |
| 20t3 | -26.6 | -27.9 | -25.3 | -25.3 | -26.9 | -27.9 | -27.7 | -25.1 | -27.4 | -27.6 | -27.5 | -25.9 | -26.2 | -25.7 | -25.4 | -24.5 |
| 20t5 | -25.8 | -27.3 | -25.3 | -24.7 | -27.1 | -27.3 | -28.2 | -25.4 | -27.0 | -27.5 | -27.7 | -25.7 | -26.0 | -25.8 | -25.0 | -24.6 |
| 21 | -25.2 | -24.4 | -25.6 | -23.3 | -24.4 | -26.4 | -26.9 | -26.6 | -27.1 | -27.6 | -26.5 | -25.5 | -26.3 | -25.6 | -24.5 | low[b] |
| 22 | -25.5 | -24.3 | -25.8 | -23.4 | -23.9 | -25.5 | -26.1 | -25.9 | -26.2 | -26.1 | -25.7 | -25.8 | -25.5 | -25.8 | -25.0 | -25.5 |
| 23 | -26.9 | -26.2 | -26.4 | -25.3 | -24.9 | -26.5 | -27.6 | -27.6 | -27.8 | -28.3 | -27.3 | -27.8 | -27.7 | -26.8 | -26.6 | -27.3 |

[a] Average values based on triplicate analysis of each sample (range in standard deviation of 0.02-1.01, with a mean standard deviation of 0.18 ‰).

[b] Concentration of the designated analyte was too low for isotope ratio measurement.

262

# Appendix C

# Appendices for Chapter 7

## C.1   Diagnostic Ratios

## C.2   Individual peak identities used in statistical dataset

Table C.1: Values for the diagnostic ratios used in Figure 7.11.

| SAMPLE | FLT/PYR | DBF/FLU | ACY/ACE |
|--------|---------|---------|---------|
| 1 | 0.84 | 0.47 | 14.10 |
| 2 | 0.58 | 0.11 | 17.12 |
| 3 | 0.87 | 0.48 | 12.46 |
| 4 | 0.90 | 0.54 | 13.32 |
| 5 | 0.93 | 0.48 | 12.31 |
| 6 | 1.04 | 0.56 | 11.11 |
| 7 | 1.20 | 1.79 | 13.52 |
| 8 | 1.17 | 1.32 | 18.19 |
| 9 | 1.33 | 1.30 | 7.38 |
| 10 | 1.21 | 1.69 | 9.74 |
| 11 | 1.94 | 1.71 | 0.11 |
| 12 | 1.62 | 1.47 | 0.22 |
| 13 | 0.82 | 1.26 | 2.69 |
| 14 | 0.86 | 0.36 | 9.41 |
| 15 | 1.00 | 0.91 | 7.39 |
| 16 | 1.22 | 1.41 | 16.31 |
| 17 | 1.19 | 1.30 | 7.28 |
| 18 | 1.25 | 1.88 | 38.24 |
| 19 | 1.47 | 1.42 | 14.33 |
| 20 T1 | 1.10 | 1.71 | 1.69 |
| 20 T3 | 1.03 | 1.68 | 1.35 |
| 20 T5 | 1.04 | 1.72 | 1.38 |
| 21 | 1.15 | 1.84 | 1.32 |
| 22 | 0.67 | 1.49 | 6.51 |
| 23 | 0.79 | 0.59 | 10.77 |
| P | 0.55 | 0.05 | 0.32 |
| D | 0.10 | 0.38 | 0.03 |
| PD | 0.08 | 0.37 | 0.09 |
| W19 | 1.43 | 1.00 | 26.25 |

Table C.2: Dataset used for statistical classification of coal tars. *(#) represents the compounds included in Figure 7.17.

| Compound Name | Target ion | No. of isomers identified | Chemical Class |
|---|---|---|---|
| undecane ($C_{11}$) * | 156 | - | (i) *n*-alkanes |
| dodecane ($C_{12}$) * | 170 | - | |
| tridecane ($C_{13}$) * | 184 | - | |
| tetradecane ($C_{14}$) * | 198 | - | |
| pentadecane ($C_{15}$) * | 212 | - | |
| hexadecane ($C_{16}$) * | 226 | - | |
| heptadecane ($C_{17}$) * | 240 | - | |
| octadecane ($C_{18}$) * | 254 | - | |
| tetracosane ($C_{24}$) | 338 | - | |
| pentacosane ($C_{25}$) | 352 | - | |
| hexacosane ($C_{26}$) | 366 | - | |
| heptacosane ($C_{27}$) * | 380 | - | |
| octacosane ($C_{28}$) * | 394 | - | |
| nonacosane ($C_{29}$) * | 408 | - | |
| triacontane ($C_{30}$) * | 422 | - | |
| hentriacontane ($C_{31}$) * | 436 | - | |
| dotriacontane ($C_{32}$) * | 450 | - | |
| tritriacontane ($C_{33}$) * | 464 | - | |
| branched alkane A ($C_{12}$) * | 170 | - | (ii) *iso*-alkanes |
| branched alkane B ($C_{13}$) * | 184 | - | |
| branched alkane C ($C_{14}$) * | 198 | - | |
| pristane * | 268 | - | |
| phytane * | 282 | - | |
| toluene * | 91 | - | (iii) alkyl benzenes |
| C3 alkyl benzene * | 120 | 3 | |
| C4 alkyl benzene * | 134 | 7 | |
| phenol * | 94 | - | (iv) phenols |
| C1 alkyl phenol * | 108 | 2 | |
| C2 alkyl phenol * | 122 | 4 | |
| 1,2-dihydronaphthalene * | 130 | - | (v) hydronaphthalenes |
| 1,4-dihydronaphthalene * | 130 | - | |
| tetrahydronaphthalene * | 132 | - | |
| naphthalene * | 128 | - | (vi) naphthalenes |
| 2-methyl naphthalene * | 142 | - | |
| 1-methyl naphthalene * | 142 | - | |
| 2-EtN | 156 | - | |
| 2,6-DMN & 2,7-DMN | 156 | - | |
| 1-EtN | 156 | - | |
| 1,6-DMN * | 156 | - | |
| 1,7-DMN & 1,3-DMN | 156 | - | |
| 2,3-DMN | 156 | - | |
| 1,4-DMN | 156 | - | |
| 1,5-DMN | 156 | - | |
| 1,2-DMN | 156 | - | |
| 1,8-DMN | 156 | - | |
| C3 alkyl naphthalene * [1] | 170 | 9 | |

Table C.3: Dataset used for statistical classification of coal tars (continued). *(#) represents the compounds included in Figure 7.17.

| Compound Name | Target ion | No. of isomers identified | Chemical Class |
|---|---|---|---|
| biphenyl | 154 | - | (vii) parent PAHs |
| acenaphthylene * | 152 | - | |
| acenaphthene * | 154 | - | |
| fluorine * | 166 | - | |
| phenanthrene * | 178 | - | |
| anthracene * | 178 | - | |
| fluoranthene * | 202 | - | |
| pyrene * | 202 | - | |
| benz[a]anthracene * | 228 | - | |
| chrysene * | 228 | - | |
| benzo[b]fluoranthene * | 252 | - | |
| benzo[k]fluoranthene * | 252 | - | |
| benzofluoranthene isomers * | 252 | 2 | |
| benzo[a]pyrene * | 252 | - | |
| benzopyrene isomer * | 252 | 1 | |
| indeno[1,2,3-cd]pyrene * | 276 | - | |
| dibenzo[a,h]anthracene * | 278 | - | |
| benzo[g,h,i]perylene * | 276 | - | |
| dibenzochrysene * | 276 | 1 | |
| phenylnaphthalene * | 204 | 1 | |
| indane * | 118 | - | |
| indene * | 116 | - | |
| diphenylmethane * | 167 | - | |
| C1 alkyl biphenyl * | 168 | 2 | (viii) alkyl PAHs |
| C1 alkyl fluorine * | 180 | 2 | |
| C1 alkyl phenanthrene/anthracene * (1) | 192 | 6 | |
| C2 alkyl phenanthrene/anthracene | 206 | 2 | |
| C1 alkyl fluoranthene/pyrene * (1) | 216 | 6 | |
| ethenyl anthracene * | 204 | 1 | |
| C1 alkyl chrysene * (1) | 242 | 3 | |
| carbazole * | 167 | - | (ix) nitrogen-containing PAHs (N-PAHs) |
| benzofuran * | 118 | - | (x) oxygen-containing PAHs (O-PAHs) |
| dibenzofuran * | 168 | - | |
| C1 alkyl dibenzofuran * | 182 | 3 | |
| benzothiophene (BT0) * | 134 | - | (xi) sulphur-containing PAHs (S-PAHs) |
| naphtho-thiophene * | 184 | - | |
| dibenzothiophene * | 184 | - | |
| phenalenothiophene * | 208 | 2 | |
| C1 alkyl benzothiophene (BT1) * | 148 | 3 | (xii) alkyl S-PAHs |
| C2 alkyl benzothiophene (BT2) * (1) | 162 | 9 | |
| C3 alkyl benzothiophene (BT3) * (1) | 176 | 5 | |
| C1 alkyl dibenzothiophene * | 198 | 3 | |
| C2 alkyl dibenzothiophene * (1) | 212 | 6 | |
| C3 alkyl dibenzothiophene * (1) | 226 | 8 | |

# C.3 Raw GCxGC data used for statistical analyses

Table C.4: Normalised GCxGC peak areas for all coal tar samples - raw data 1/10

| Compound | Mass (g/mol) | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| undecane (C11) | 156 | 1.793E-01 | 1.464E-01 | 1.137E-01 | 3.017E-01 | 4.093E-03 | 1.830E-03 | 2.844E-02 | 4.745E-01 | 2.243E+00 | 2.598E-01 | 2.466E-01 | 4.875E-01 |
| dodecane (C12) | 170 | 1.800E-01 | 1.340E-01 | 1.601E-01 | 3.758E-01 | 1.589E-02 | 2.905E-03 | 1.897E-02 | 6.192E-01 | 1.730E+00 | 4.149E-01 | 3.502E-01 | 5.989E-01 |
| tridecane (C13) | 184 | 2.982E-01 | 1.333E-01 | 1.706E-01 | 4.040E-01 | 3.139E-02 | 6.976E-03 | 8.596E-03 | 1.005E+00 | 1.089E+00 | 2.922E-01 | 3.364E-01 | 2.432E-07 |
| tetradecane (C14) | 198 | 1.583E-01 | 6.639E-02 | 9.102E-02 | 2.905E-01 | 3.201E-02 | 6.659E-03 | 9.694E-03 | 4.475E-01 | 8.751E-01 | 7.110E-02 | 2.487E+01 | 2.432E-07 |
| pentadecane (C15) | 212 | 7.927E-02 | 3.939E-02 | 4.505E-02 | 2.175E-01 | 1.927E-02 | 7.172E-05 | 1.779E-05 | 4.249E-01 | 3.558E-01 | 2.639E-02 | 1.635E+01 | 2.432E-07 |
| hexadecane (C16) | 226 | 2.459E-02 | 3.558E-02 | 4.624E-02 | 1.269E-01 | 7.423E-03 | 2.169E-01 | 1.779E-05 | 1.410E-01 | 1.499E-01 | 1.739E-05 | 1.092E-01 | 2.432E-07 |
| heptadecane (C17) | 240 | 2.549E-07 | 1.136E-02 | 2.269E-02 | 1.189E-01 | 3.622E-03 | 1.707E-03 | 2.539E-07 | 1.178E-01 | 9.683E-02 | 2.491E-07 | 8.454E+00 | 2.432E-07 |
| octadecane (C18) | 254 | 2.549E-07 | 1.311E-03 | 1.307E-02 | 4.503E-02 | 2.836E-02 | 2.562E-07 | 2.539E-07 | 5.023E-02 | 5.487E-02 | 2.491E-07 | 5.628E+00 | 2.432E-07 |
| tetracosane (C24) | 338 | 2.549E-07 | 2.594E-07 | 2.533E-07 | 8.298E-03 | 2.536E-07 | 2.562E-07 | 2.539E-07 | 2.558E-07 | 3.291E-02 | 2.491E-07 | 5.666E-01 | 2.432E-07 |
| pentacosane (C25) | 352 | 2.549E-07 | 2.594E-07 | 2.533E-07 | 6.420E-03 | 2.536E-07 | 2.562E-07 | 2.539E-07 | 2.558E-07 | 5.767E-02 | 2.491E-07 | 3.596E-01 | 2.432E-07 |
| hexacosane (C26) | 366 | 2.549E-07 | 2.594E-07 | 2.533E-07 | 8.937E-03 | 2.536E-07 | 2.562E-07 | 2.539E-07 | 2.558E-07 | 3.995E-02 | 2.491E-07 | 1.986E-01 | 2.432E-07 |
| heptacosane (C27) | 380 | 2.549E-07 | 2.594E-07 | 2.533E-07 | 5.280E-03 | 2.536E-07 | 2.562E-07 | 2.539E-07 | 2.558E-07 | 5.402E-02 | 2.491E-07 | 1.401E-01 | 2.432E-07 |
| octacosane (C28) | 394 | 2.549E-07 | 2.594E-07 | 2.533E-07 | 2.943E-03 | 2.536E-07 | 2.562E-07 | 2.539E-07 | 2.558E-07 | 4.514E-02 | 2.491E-07 | 4.529E-02 | 2.432E-07 |
| nonacosane (C29) | 408 | 2.549E-07 | 2.594E-07 | 2.533E-07 | 3.083E-03 | 2.536E-07 | 2.562E-07 | 2.539E-07 | 2.558E-07 | 2.060E-02 | 2.491E-07 | 3.840E-02 | 2.432E-07 |
| triacontane (C30) | 422 | 2.549E-07 | 2.594E-07 | 2.533E-07 | 2.519E-07 | 2.536E-07 | 2.562E-07 | 2.539E-07 | 2.558E-07 | 2.501E-07 | 2.491E-07 | 2.825E-02 | 2.432E-07 |
| hentriacontane (C31) | 436 | 2.549E-07 | 2.594E-07 | 2.533E-07 | 2.519E-07 | 2.536E-07 | 2.562E-07 | 2.539E-07 | 2.558E-07 | 2.501E-07 | 2.491E-07 | 1.883E-02 | 2.432E-07 |
| dotriacontane (C32) | 450 | 2.549E-07 | 2.594E-07 | 2.533E-07 | 2.519E-07 | 2.536E-07 | 2.562E-07 | 2.539E-07 | 2.558E-07 | 2.501E-07 | 2.491E-07 | 2.470E-07 | 2.432E-07 |
| tritriacontane (C33) | 464 | 2.549E-07 | 2.594E-07 | 2.533E-07 | 2.519E-07 | 2.536E-07 | 2.562E-07 | 2.539E-07 | 2.558E-07 | 2.501E-07 | 2.491E-07 | 2.470E-07 | 2.432E-07 |
| pristane | 170 | 1.429E-01 | 2.388E-02 | 8.844E-02 | 1.659E-01 | 2.806E-02 | 1.378E-02 | 2.539E-07 | 2.772E-01 | 4.953E-01 | 3.465E-02 | 5.857E+00 | 1.402E-01 |
| phytane | 184 | 2.549E-07 | 1.574E-02 | 2.103E-02 | 1.469E-02 | 2.573E-03 | 3.140E-05 | 2.539E-07 | 2.515E-02 | 4.252E-07 | 2.491E-07 | 1.926E+00 | 1.792E-01 |
| branched alkane A (C12) | 198 | 1.302E-01 | 4.210E-02 | 8.801E-02 | 1.580E-01 | 1.581E-02 | 5.911E-06 | 2.285E-02 | 5.275E-01 | 1.038E+00 | 1.935E-01 | 1.439E-01 | 2.843E-01 |
| branched alkane B (C13) | 268 | 1.479E-01 | 3.632E-02 | 9.022E-02 | 1.233E-01 | 1.991E-02 | 1.002E-02 | 1.382E-02 | 3.184E-01 | 4.102E-01 | 1.297E-01 | 7.723E+00 | 3.849E-01 |
| branched alkane C (C14) | 282 | 1.711E-01 | 3.372E-02 | 7.550E-02 | 1.129E-01 | 2.220E-02 | 9.404E-03 | 5.246E-04 | 2.938E-01 | 3.240E-01 | 8.717E-02 | 6.848E+00 | 3.703E+01 |
| phenol | 94 | 2.549E-07 | 2.594E-07 | 2.533E-07 | 2.519E-07 | 2.536E-07 | 2.562E-07 | 3.966E+00 | 3.754E-01 | 3.448E-01 | 1.564E+00 | 2.470E-07 | 2.432E-07 |
| C1 phenol a | 108 | 2.549E-07 | 2.594E-07 | 2.533E-07 | 3.841E-02 | 2.536E-07 | 2.562E-07 | 2.034E+00 | 5.581E-01 | 1.228E+00 | 2.426E+00 | 3.036E-02 | 2.432E-07 |
| C1 phenol b | 108 | 2.549E-07 | 2.261E-03 | 2.533E-07 | 4.790E-03 | 2.536E-07 | 2.562E-07 | 3.265E+00 | 7.303E-01 | 8.030E-01 | 2.032E+00 | 3.802E-03 | 2.432E-07 |
| C2 phenol a | 122 | 2.549E-07 | 2.819E-03 | 2.533E-07 | 1.310E-02 | 2.536E-07 | 2.562E-07 | 2.875E-01 | 2.325E-01 | 4.825E-01 | 7.982E-01 | 1.885E-01 | 2.432E-07 |
| C2 phenol b | 122 | 2.549E-07 | 2.594E-07 | 2.533E-07 | 6.755E-02 | 2.536E-07 | 2.562E-07 | 1.621E+00 | 1.110E+00 | 2.384E+00 | 1.806E+00 | 6.019E-01 | 2.432E-07 |
| C2 phenol c | 122 | 2.549E-07 | 2.594E-07 | 2.533E-07 | 9.677E-03 | 2.536E-07 | 2.562E-07 | 5.946E-01 | 2.128E-01 | 5.436E-01 | 3.829E-01 | 9.561E-02 | 2.432E-07 |
| C2 phenol d | 122 | 2.549E-07 | 2.594E-07 | 2.533E-07 | 8.615E-03 | 2.536E-07 | 2.562E-07 | 2.082E-01 | 1.220E-01 | 1.239E-01 | 1.979E-01 | 6.069E-02 | 2.432E-07 |
| toluene | 91 | 1.707E-01 | 1.021E+01 | 1.204E+00 | 6.166E+00 | 4.541E+00 | 5.068E-01 | 1.728E+01 | 2.558E-01 | 3.134E+01 | 2.262E+01 | 5.995E+00 | 2.795E-02 |
| C3-enyl benzene | 120 | 2.928E+00 | 6.517E+00 | 8.119E-01 | 2.493E+00 | 6.397E-02 | 5.954E-02 | 9.671E-01 | 7.137E-01 | 4.482E-01 | 3.486E+00 | 2.678E+00 | 3.504E+00 |
| C3 benzene a | 120 | 2.660E+00 | 4.387E+00 | 1.138E+00 | 2.459E+00 | 3.090E-02 | 5.168E-02 | 3.102E+00 | 8.511E+01 | 1.331E+01 | 9.134E+00 | 1.907E+01 | 8.176E+00 |
| C3 benzene b | 120 | 9.205E-01 | 1.350E+00 | 4.437E-01 | 7.625E-01 | 1.842E-02 | 1.991E-02 | 9.817E-01 | 4.025E+00 | 3.839E+00 | 2.619E+00 | 7.461E+00 | 3.402E+00 |
| C4 benzene a | 134 | 1.421E-01 | 1.004E-01 | 3.884E-02 | 1.577E-01 | 1.660E-03 | 1.056E-01 | 5.381E-02 | 2.813E-01 | 1.127E+00 | 4.417E-01 | 3.638E+00 | 3.828E+00 |
| C4 benzene b | 134 | 2.891E-01 | 3.900E-01 | 1.581E-01 | 3.389E-01 | 8.591E-03 | 8.034E-01 | 2.335E-01 | 1.107E+00 | 2.112E+00 | 1.025E+00 | 9.213E+00 | 6.045E+00 |
| C4 benzene c | 134 | 4.829E-01 | 5.058E-01 | 1.692E-01 | 4.332E-01 | 1.429E-02 | 1.353E-01 | 1.393E-01 | 1.669E+00 | 2.054E+00 | 9.270E-01 | 1.418E-01 | 1.026E+00 |
| C4 benzene d | 134 | 5.004E-01 | 4.886E-01 | 2.310E-01 | 4.971E-01 | 1.729E-02 | 1.727E-02 | 1.231E-01 | 1.711E+00 | 2.501E+00 | 1.043E+00 | 1.649E-01 | 1.744E-01 |
| C4 benzene e | 134 | 4.415E-01 | 6.639E-01 | 2.291E-01 | 4.836E-01 | 2.085E-02 | 9.678E-03 | 3.103E-01 | 3.164E+00 | 1.458E+00 | 1.090E+00 | 9.204E+00 | 1.195E-01 |
| C4 benzene f | 134 | 5.836E-01 | 7.499E-01 | 2.696E-01 | 4.566E-01 | 2.990E-02 | 1.583E-02 | 3.264E-01 | 1.490E+00 | 1.112E+00 | 1.052E+00 | 1.605E-01 | 1.480E-01 |
| C4 benzene g | 134 | 6.612E-01 | 8.440E-01 | 2.931E-01 | 5.330E-01 | 3.753E-02 | 2.196E-02 | 3.745E-01 | 2.696E+00 | 1.378E+00 | 9.645E-01 | 1.122E-01 | 2.016E-01 |

Table C.5: Normalised GCxGC peak areas for all coal tar samples - raw data 2/10

| Compound | Mass (g/mol) | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| indane | 118 | 7.661E-01 | 1.177E+00 | 2.080E-01 | 5.387E-01 | 1.036E-02 | 2.324E-02 | 6.375E-01 | 2.298E+00 | 3.655E-01 | 2.187E+00 | 2.436E+01 | 6.267E+00 |
| indene | 116 | 1.721E-01 | 1.965E+01 | 4.505E+00 | 1.206E+01 | 4.602E-01 | 4.698E-01 | 1.839E+01 | 1.101E+02 | 3.252E+01 | 8.857E+01 | 4.357E+01 | 2.873E+00 |
| benzofuran | 118 | 4.402E-01 | 7.479E-01 | 5.577E-02 | 3.073E-01 | 4.166E-02 | 8.445E-03 | 6.064E+00 | 8.539E+00 | 5.679E+00 | 1.701E+01 | 7.879E+00 | 6.883E-02 |
| benzothiophene | 134 | 1.102E+01 | 1.122E-01 | 3.931E+00 | 7.287E+00 | 9.866E-01 | 5.752E-01 | 2.623E-01 | 1.074E+02 | 3.284E-01 | 7.859E-01 | 5.595E-01 | 1.080E+01 |
| C1 benzothiophene a | 148 | 1.994E+00 | 4.637E+00 | 1.605E+00 | 2.915E+00 | 3.080E-01 | 1.106E-01 | 1.050E+00 | 1.053E-01 | 2.040E+00 | 3.144E+00 | 1.111E+01 | 9.176E+00 |
| C1 benzothiophene b | 148 | 1.096E+00 | 1.523E+00 | 7.143E-01 | 1.071E+00 | 2.492E-01 | 6.453E-02 | 8.343E-01 | 5.643E+00 | 1.171E+00 | 1.541E+00 | 8.697E+00 | 1.570E+00 |
| C1 benzothiophene c | 148 | 2.808E+00 | 4.630E+00 | 1.205E+00 | 1.769E+00 | 3.561E-01 | 1.430E-01 | 1.115E+00 | 1.257E-01 | 1.703E+00 | 1.736E+00 | 9.088E+00 | 1.005E+01 |
| C2 benzothiophene a | 162 | 3.322E+00 | 6.196E-01 | 2.193E-01 | 2.969E+00 | 4.060E-02 | 1.856E-02 | 2.040E-02 | 1.522E+00 | 2.402E-01 | 1.277E-01 | 1.995E+00 | 3.979E+00 |
| C2 benzothiophene b | 162 | 8.785E-01 | 1.505E-01 | 5.814E-02 | 7.157E-02 | 1.831E-02 | 8.318E-03 | 2.759E-02 | 3.755E-01 | 1.518E-01 | 9.730E-02 | 1.116E+00 | 3.082E+00 |
| C2 benzothiophene c | 162 | 1.779E-01 | 5.381E-01 | 2.486E-01 | 3.474E-01 | 4.828E-02 | 1.879E-02 | 1.174E-01 | 1.684E+00 | 2.061E-01 | 1.059E-01 | 1.812E+00 | 3.308E+00 |
| C2 benzothiophene d | 162 | 1.821E+00 | 8.308E-01 | 2.210E-01 | 2.726E-01 | 4.954E-02 | 2.373E-02 | 1.026E-01 | 1.582E+00 | 4.135E-01 | 2.113E-01 | 3.012E+00 | 5.067E+00 |
| C2 benzothiophene e | 162 | 5.481E-01 | 1.115E+00 | 2.526E-01 | 3.475E-01 | 6.847E-02 | 3.659E-02 | 1.549E-01 | 2.316E+00 | 3.924E-01 | 1.742E-01 | 2.482E+00 | 4.051E+00 |
| C2 benzothiophene f | 162 | 3.291E-01 | 5.464E-01 | 2.230E-01 | 2.985E-01 | 5.639E-02 | 2.373E-02 | 1.042E-01 | 1.603E+00 | 1.503E-01 | 1.426E-01 | 2.097E+00 | 3.096E+00 |
| C2 benzothiophene g | 162 | 2.687E-01 | 2.219E-01 | 1.498E-01 | 1.783E-01 | 3.662E-02 | 1.601E-02 | 5.771E-02 | 1.199E+00 | 2.330E-01 | 6.291E-02 | 1.403E+00 | 6.214E+00 |
| C2 benzothiophene h | 162 | 4.097E-02 | 3.698E-02 | 2.031E-02 | 3.838E-02 | 9.563E-03 | 4.548E-03 | 2.116E-02 | 2.023E-01 | 3.503E-02 | 1.685E-02 | 2.655E-01 | 4.586E-01 |
| C2 benzothiophene i | 162 | 8.154E-03 | 2.379E-02 | 7.684E-03 | 1.323E-02 | 1.019E-03 | 8.132E-04 | 6.489E-03 | 9.460E-02 | 1.554E-02 | 9.471E-03 | 2.220E-01 | 3.103E-01 |
| C3 benzothiophene a | 176 | 3.226E-02 | 5.039E-02 | 3.259E-02 | 3.374E-02 | 5.363E-03 | 2.743E-03 | 2.355E-02 | 1.915E-01 | 6.990E-02 | 2.058E-02 | 5.511E-01 | 7.865E-01 |
| C3 benzothiophene b | 176 | 1.223E-01 | 2.739E-01 | 7.798E-02 | 1.430E-01 | 2.160E-02 | 9.422E-03 | 2.276E-02 | 8.559E-01 | 8.712E-01 | 4.636E-02 | 1.035E+00 | 4.361E+00 |
| C3 benzothiophene c | 176 | 5.365E-02 | 8.299E-02 | 2.403E-02 | 8.121E-02 | 7.844E-03 | 3.903E-03 | 1.071E-02 | 1.713E-01 | 5.936E-02 | 2.327E-02 | 6.327E-01 | 2.345E+00 |
| C3 benzothiophene d | 176 | 1.400E-01 | 2.789E-01 | 1.275E-01 | 1.702E-01 | 2.826E-02 | 8.228E-03 | 1.467E-02 | 7.292E-01 | 1.263E-01 | 3.126E-02 | 2.840E+00 | 8.386E+00 |
| C3 benzothiophene e | 176 | 3.580E-02 | 8.086E-02 | 4.500E-02 | 8.949E-02 | 1.183E-02 | 3.340E-03 | 2.823E-02 | 4.039E-01 | 5.657E-02 | 2.002E-02 | 9.228E-01 | 5.121E+00 |
| diphenylmethane | 167 | 4.700E-02 | 2.483E-01 | 8.135E-02 | 1.142E-01 | 1.069E-02 | 2.523E-03 | 4.249E-02 | 3.212E-01 | 5.176E-02 | 3.937E-02 | 9.774E-01 | 1.713E+00 |
| 1,2-dihydronaphthalene | 130 | 2.012E+00 | 5.143E+00 | 9.556E-01 | 1.932E+00 | 1.280E-01 | 6.654E-02 | 1.243E+00 | 1.258E-01 | 2.715E+00 | 4.128E+00 | 7.257E+00 | 2.074E+00 |
| 1,4-dihydronaphthalene | 130 | 3.139E+00 | 4.928E+00 | 1.271E+00 | 1.987E+00 | 2.289E-01 | 1.387E-01 | 8.703E-01 | 1.011E-01 | 2.471E+00 | 4.650E+00 | 7.184E+00 | 1.757E+00 |
| tetrahydronaphthalene | 132 | 8.518E-02 | 1.121E-01 | 3.161E-02 | 6.334E-02 | 3.803E-03 | 3.255E-03 | 2.155E-02 | 1.520E-01 | 1.832E-01 | 1.304E-01 | 3.303E+00 | 1.126E+00 |
| naphthalene | 128 | 2.538E-02 | 1.970E+02 | 7.535E+01 | 1.202E+02 | 1.693E+01 | 9.125E+00 | 5.354E+02 | 1.568E+03 | 5.204E+02 | 1.442E+03 | 5.182E+02 | 6.933E+01 |
| 2-methyl naphthalene | 142 | 3.785E-01 | 4.677E+01 | 1.659E+01 | 2.553E+01 | 3.871E+00 | 1.798E+00 | 2.615E+01 | 1.615E+02 | 2.957E+01 | 5.199E+01 | 2.260E+02 | 3.243E+01 |
| 1-methyl naphthalene | 142 | 2.624E-01 | 3.426E+01 | 1.368E+01 | 1.813E+01 | 2.979E+00 | 1.424E+00 | 1.513E+01 | 1.017E+02 | 1.806E+01 | 2.684E+01 | 9.996E+01 | 6.364E+01 |
| 2-EtN | 156 | 1.704E+00 | 3.762E+00 | 2.325E+00 | 3.702E+00 | 3.572E-01 | 9.846E-02 | 8.370E-01 | 1.137E-01 | 2.038E+00 | 1.370E+00 | 3.641E+01 | 3.037E+01 |
| 2,6-DMN & 2,7-DMN | 156 | 4.477E+00 | 5.850E+00 | 4.074E+00 | 5.130E+00 | 9.281E-01 | 4.077E-01 | 3.815E+00 | 2.482E+01 | 6.248E+00 | 5.820E+00 | 9.201E+01 | 1.007E+02 |
| 1-EtN | 156 | 4.801E+00 | 1.836E+00 | 1.466E+00 | 1.462E+00 | 6.236E-01 | 3.968E-01 | 2.236E+00 | 8.063E+00 | 4.129E+00 | 3.637E+00 | 1.494E+01 | 9.956E+00 |
| 1,6-DMN | 156 | 7.380E+00 | 9.323E+00 | 6.087E+00 | 8.430E+00 | 1.383E+00 | 5.962E+00 | 3.921E+00 | 3.370E+01 | 8.278E+00 | 6.048E+00 | 1.837E+02 | 9.307E+01 |
| 1,7-DMN & 1,3-DMN | 156 | 1.573E+00 | 1.199E+01 | 7.154E+00 | 9.024E+00 | 1.996E+00 | 1.320E+00 | 8.161E+00 | 5.783E+00 | 1.745E+00 | 1.248E+00 | 8.005E+01 | 6.833E+01 |
| 2,3-DMN | 156 | 2.275E+00 | 2.855E+00 | 1.962E+00 | 2.756E+00 | 4.247E-01 | 2.124E-01 | 1.776E+00 | 1.014E+01 | 3.744E+00 | 2.420E+00 | 3.467E+01 | 4.013E+01 |
| 1,4-DMN | 156 | 2.151E+00 | 2.834E+00 | 1.583E+00 | 2.097E+00 | 4.827E-01 | 2.220E-01 | 1.328E+00 | 8.908E+00 | 2.597E+00 | 1.910E+00 | 1.640E+01 | 2.412E+01 |
| 1,5-DMN | 156 | 2.944E+00 | 2.157E+00 | 1.290E+00 | 1.812E+00 | 4.073E-01 | 2.175E-01 | 1.682E+00 | 9.838E+00 | 3.595E+00 | 2.138E+00 | 1.623E+01 | 2.316E+01 |
| 1,2-DMN | 156 | 2.562E+00 | 1.946E+00 | 1.318E+00 | 1.794E+00 | 4.918E-01 | 2.468E-01 | 1.516E+00 | 1.034E+01 | 1.608E+00 | 2.055E+00 | 2.156E+01 | 2.996E+01 |
| 1,8-DMN | 156 | 1.146E-01 | 1.821E-01 | 3.360E-02 | 4.541E-02 | 3.151E-03 | 2.287E-03 | 2.539E-07 | 2.558E-07 | 2.501E-07 | 2.491E-07 | 2.407E-07 | 1.261E+00 |
| C3 alkyl naphthalene a | 170 | 5.171E-02 | 1.459E-01 | 1.108E-01 | 1.873E-01 | 1.811E-02 | 8.796E-03 | 8.573E-02 | 3.146E-01 | 1.701E-01 | 4.240E-02 | 3.495E-01 | 1.431E+00 |
| C3 alkyl naphthalene b | 170 | 2.337E-01 | 1.683E-01 | 1.704E-01 | 4.202E-01 | 5.019E-02 | 3.822E-02 | 9.555E-02 | 1.519E+00 | 5.806E-01 | 1.984E+00 | 9.732E-01 | 1.099E+01 |

Table C.6: Normalised GCxGC peak areas for all coal tar samples - raw data 3/10

| Compound | Mass (g/mol) | Samples | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
| C3 alkyl naphthalene c | 170 | 1.813E-01 | 7.340E-02 | 2.653E-01 | 1.238E-01 | 1.985E-02 | 1.586E-02 | 5.109E-02 | 6.621E-01 | 3.382E-01 | 8.821E-02 | 2.792E+00 | 2.803E+00 |
| C3 alkyl naphthalene d | 170 | 1.443E+00 | 8.781E-01 | 8.785E-01 | 1.093E+00 | 2.095E-01 | 1.114E-01 | 4.342E-01 | 4.341E-01 | 2.165E+00 | 1.090E+00 | 1.342E-01 | 2.224E+01 |
| C3 alkyl naphthalene e | 170 | 8.848E-01 | 6.124E-01 | 6.405E-01 | 8.024E-01 | 1.821E-01 | 1.028E-01 | 5.687E-01 | 3.609E+00 | 1.724E+00 | 8.793E-01 | 1.222E+01 | 1.761E+01 |
| C3 alkyl naphthalene f | 170 | 5.100E-01 | 2.601E-01 | 2.675E-01 | 4.088E-01 | 9.893E-02 | 3.683E-02 | 2.628E-01 | 1.543E+00 | 7.661E-01 | 3.381E-01 | 6.377E+00 | 8.836E+00 |
| C3 alkyl naphthalene g | 170 | 3.395E-01 | 1.658E-01 | 1.947E-01 | 2.556E-01 | 6.543E-02 | 4.072E-02 | 1.516E-01 | 1.217E+00 | 5.202E-01 | 2.140E-01 | 2.082E+00 | 3.954E+00 |
| C3 alkyl naphthalene h | 170 | 6.843E-02 | 5.601E-02 | 5.105E-02 | 9.935E-02 | 1.101E-02 | 6.413E-03 | 1.867E-02 | 2.493E-01 | 8.561E-02 | 2.533E-02 | 8.267E-01 | 1.336E+00 |
| C3 alkyl naphthalene i | 170 | 2.904E-01 | 8.584E-02 | 1.619E-01 | 1.782E-01 | 3.018E-02 | 2.085E-02 | 1.078E-01 | 7.065E-01 | 7.480E-01 | 1.780E-01 | 4.316E+00 | 4.286E+00 |
| biphenyl | 154 | 3.062E+00 | 7.073E+00 | 1.822E+00 | 2.568E+00 | 4.252E-01 | 1.333E-01 | 5.672E+00 | 1.980E+00 | 4.494E+00 | 9.504E+00 | 4.310E+01 | 4.330E+01 |
| C1 biphenyl a | 168 | 5.102E-01 | 1.054E+00 | 5.025E-01 | 6.715E-01 | 1.296E-01 | 3.886E-02 | 4.753E-01 | 3.702E+00 | 5.150E-01 | 7.945E-01 | 8.145E+01 | 1.087E+01 |
| C1 biphenyl b | 168 | 1.931E-01 | 4.968E-01 | 2.012E-01 | 3.556E-01 | 2.448E-02 | 1.386E-02 | 2.226E-01 | 1.682E+00 | 2.171E-01 | 2.658E-01 | 2.865E+00 | 5.155E+00 |
| dibenzofuran | 168 | 2.916E+00 | 6.812E-01 | 1.613E+00 | 2.382E+00 | 4.791E-01 | 2.899E-01 | 2.460E+01 | 5.318E+01 | 1.753E+01 | 3.455E+01 | 1.246E+02 | 6.674E+01 |
| C1 dibenzofuran a | 182 | 4.725E-01 | 1.907E-01 | 3.157E-01 | 4.310E-01 | 7.701E-02 | 5.048E-02 | 2.060E+00 | 3.969E+00 | 2.076E+00 | 2.399E+00 | 1.081E+01 | 1.760E+01 |
| C1 dibenzofuran b | 182 | 3.615E-01 | 1.748E-01 | 3.658E-01 | 5.224E-01 | 8.998E-02 | 5.030E-02 | 2.663E+00 | 5.301E+00 | 2.895E+00 | 2.950E+00 | 1.405E+01 | 2.502E+01 |
| C1 dibenzofuran c | 182 | 5.017E-02 | 6.531E-02 | 9.565E-02 | 1.459E-01 | 1.594E-02 | 1.418E-02 | 4.645E-01 | 6.337E-01 | 6.853E-01 | 5.376E-01 | 2.578E+00 | 3.995E+00 |
| acenaphthylene | 152 | 1.860E-01 | 2.436E+01 | 8.572E-01 | 1.238E-01 | 2.580E+00 | 1.328E+00 | 2.927E-01 | 1.167E-02 | 2.501E-01 | 3.823E-01 | 1.088E-01 | 4.979E+00 |
| acenaphthene | 154 | 1.319E+00 | 1.424E+00 | 6.866E-01 | 9.302E-01 | 2.096E-01 | 1.197E-01 | 2.166E+00 | 6.417E+00 | 3.391E+00 | 3.926E+00 | 9.483E+01 | 2.338E+01 |
| dibenzothiophene | 184 | 2.882E+00 | 3.222E+00 | 1.790E+00 | 2.367E+00 | 4.789E-01 | 2.564E-01 | 2.789E+00 | 1.270E-01 | 3.134E+00 | 5.063E+00 | 1.810E+01 | 3.273E+01 |
| naphtho-thiophene | 184 | 7.725E-01 | 5.541E-01 | 3.667E-01 | 4.916E-01 | 1.339E-01 | 1.021E-01 | 1.825E+00 | 4.682E+00 | 2.171E+00 | 3.011E+00 | 5.269E+00 | 1.198E+01 |
| C1 dibenzothiophene a | 198 | 7.910E-01 | 1.163E+00 | 5.635E-01 | 6.779E-01 | 1.550E-01 | 7.971E-02 | 1.342E-01 | 3.356E+00 | 1.830E-01 | 2.131E-01 | 1.835E+00 | 1.054E+01 |
| C1 dibenzothiophene b | 198 | 3.428E-01 | 3.800E-01 | 2.241E-01 | 3.903E-01 | 6.882E-02 | 3.174E-02 | 8.207E-02 | 1.200E+00 | 1.320E-01 | 1.625E-01 | 8.356E-01 | 4.340E+00 |
| C1 dibenzothiophene c | 198 | 4.887E-01 | 4.378E-01 | 2.904E-01 | 3.457E-01 | 7.555E-02 | 4.498E-02 | 1.266E-01 | 1.845E+00 | 1.796E-01 | 1.931E-01 | 9.884E-01 | 5.877E+00 |
| C2 dibenzothiophene a | 212 | 3.525E-03 | 1.405E-02 | 7.293E-03 | 1.013E-02 | 8.419E-04 | 8.737E-05 | 2.539E-07 | 4.235E-02 | 2.547E-03 | 2.829E-03 | 1.182E-01 | 1.364E+00 |
| C2 dibenzothiophene b | 212 | 1.127E-01 | 2.173E-01 | 7.889E-02 | 1.086E-01 | 2.075E-02 | 1.076E-02 | 8.658E-03 | 6.456E-01 | 2.230E-02 | 2.317E-02 | 3.556E-01 | 3.815E+00 |
| C2 dibenzothiophene c | 212 | 5.050E-02 | 8.356E-02 | 3.688E-02 | 5.236E-02 | 1.361E-02 | 3.085E-03 | 9.876E-03 | 2.234E-01 | 1.056E-02 | 1.321E-02 | 1.603E-01 | 1.479E+00 |
| C2 dibenzothiophene d | 212 | 8.250E-02 | 1.415E-01 | 5.938E-02 | 6.159E-02 | 1.536E-02 | 6.316E-03 | 1.524E-02 | 3.538E-01 | 1.802E-02 | 1.611E-02 | 2.709E-01 | 1.852E+00 |
| C2 dibenzothiophene e | 212 | 7.761E-02 | 1.408E-01 | 4.997E-02 | 6.815E-02 | 1.230E-02 | 7.421E-03 | 8.367E-03 | 4.485E-01 | 2.280E-02 | 1.084E-02 | 3.482E+00 | 3.038E+00 |
| C2 dibenzothiophene f | 212 | 9.660E-02 | 5.003E-02 | 4.656E-02 | 4.097E-02 | 2.114E-02 | 9.776E-03 | 9.648E-03 | 8.738E-01 | 1.864E-02 | 4.883E-03 | 1.097E-01 | 5.129E-01 |
| C3 dibenzothiophene a | 226 | 1.638E-02 | 3.666E-02 | 1.721E-02 | 1.673E-02 | 3.162E-03 | 2.150E-03 | 2.534E-03 | 9.632E-02 | 7.050E-03 | 1.390E-02 | 9.078E-01 | 9.803E-01 |
| C3 dibenzothiophene b | 226 | 2.863E-03 | 8.617E-03 | 7.121E-03 | 8.560E-03 | 1.181E-03 | 3.408E-04 | 2.640E-03 | 4.297E-02 | 4.523E-03 | 5.745E-03 | 7.107E-02 | 5.456E-01 |
| C3 dibenzothiophene c | 226 | 1.879E-03 | 1.722E-02 | 7.817E-03 | 1.385E-02 | 1.016E-03 | 1.131E-03 | 4.546E-03 | 3.883E-02 | 6.190E-03 | 3.780E-03 | 5.849E-02 | 4.111E-01 |
| C3 dibenzothiophene d | 226 | 5.468E-03 | 1.153E-02 | 2.051E-03 | 3.095E-03 | 9.692E-04 | 6.044E-04 | 2.564E-03 | 5.704E-02 | 5.861E-03 | 1.828E-03 | 4.722E-01 | 5.768E-01 |
| C3 dibenzothiophene e | 226 | 4.791E-03 | 1.982E-02 | 1.144E-02 | 1.236E-02 | 3.063E-03 | 1.941E-03 | 6.043E-03 | 1.049E-01 | 6.857E-03 | 5.880E-03 | 7.912E-02 | 1.037E+00 |
| C3 dibenzothiophene f | 226 | 1.326E-02 | 3.128E-02 | 1.296E-02 | 1.805E-02 | 2.226E-03 | 1.728E-03 | 6.374E-03 | 1.487E-01 | 5.899E-03 | 8.279E-03 | 1.383E-01 | 8.628E-01 |
| C3 dibenzothiophene g | 226 | 6.666E-03 | 1.524E-02 | 8.423E-03 | 9.267E-03 | 1.964E-03 | 8.983E-04 | 7.670E-03 | 7.891E-02 | 3.452E-03 | 3.465E-03 | 4.811E-01 | 6.041E-01 |
| C3 dibenzothiophene h | 226 | 7.954E-03 | 1.840E-02 | 5.409E-03 | 7.590E-03 | 1.947E-03 | 1.206E-03 | 6.563E-03 | 5.507E-02 | 2.561E-03 | 2.384E-03 | 5.115E-02 | 4.472E-01 |
| fluorene | 166 | 6.192E+00 | 6.225E+00 | 3.358E+00 | 4.408E+00 | 9.962E-01 | 5.129E-01 | 1.375E-01 | 4.016E-01 | 1.345E-01 | 2.041E-01 | 7.289E+01 | 4.545E+01 |
| C1 fluorene a | 180 | 1.662E-01 | 3.716E-01 | 1.286E-01 | 1.697E-01 | 2.285E-02 | 5.672E-03 | 1.723E-01 | 6.878E-01 | 1.571E-01 | 9.240E-02 | 8.188E-02 | 4.253E+01 |
| C1 fluorene b | 180 | 3.367E-02 | 7.482E-02 | 2.240E-02 | 5.859E-02 | 1.516E-02 | 5.772E-03 | 7.416E-02 | 2.219E-01 | 6.703E-02 | 3.319E-02 | 2.419E-01 | 2.110E+01 |
| phenanthrene | 178 | 1.455E-01 | 1.461E+01 | 8.692E+00 | 1.239E+01 | 2.363E+00 | 1.527E+00 | 6.415E+01 | 1.662E+02 | 4.871E+01 | 9.509E+01 | 2.530E+02 | 1.054E+02 |
| anthracene | 178 | 7.025E+00 | 4.385E+00 | 2.935E+00 | 3.811E+00 | 1.119E+00 | 6.738E-01 | 2.080E+01 | 3.974E+01 | 2.018E+01 | 2.585E+01 | 4.678E+01 | 4.719E+01 |

Table C.7: Normalised GCxGC peak areas for all coal tar samples - raw data 4/10

| Compound | Mass (g/mol) | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| C1 phenanthrene a | 192 | 7.647E-01 | 7.216E-01 | 6.347E-01 | 8.244E-01 | 1.557E-01 | 7.555E-02 | 1.175E+00 | 4.206E+00 | 1.199E+00 | 1.644E+00 | 4.452E-01 | 1.337E+01 |
| C1 phenanthrene b | 192 | 6.069E-01 | 9.973E-01 | 6.327E-01 | 8.720E-01 | 1.326E-01 | 6.669E-02 | 1.383E+00 | 4.581E+00 | 1.149E+00 | 1.301E+00 | 5.262E+00 | 1.647E+01 |
| C1 phenanthrene c | 192 | 1.055E+00 | 5.411E-01 | 3.959E-01 | 4.443E-01 | 1.177E-01 | 1.044E-01 | 1.433E+00 | 3.359E+00 | 1.786E+00 | 1.385E+00 | 3.312E+00 | 3.523E+00 |
| C1 anthracene a | 192 | 3.018E-01 | 3.028E-01 | 1.813E-01 | 2.386E-01 | 5.393E-02 | 3.259E-02 | 4.203E-01 | 1.316E+00 | 5.461E-01 | 6.000E-01 | 1.498E+00 | 2.485E+00 |
| C1 anthracene b | 192 | 6.948E-01 | 4.874E-01 | 4.311E-01 | 5.044E-01 | 1.118E-01 | 5.137E-02 | 6.369E-01 | 2.433E+00 | 9.163E-01 | 9.529E-01 | 2.329E+00 | 6.870E+00 |
| C1 anthracene c | 192 | 1.088E+00 | 1.109E+00 | 4.877E-01 | 5.417E-01 | 1.853E-01 | 1.348E-01 | 1.016E+00 | 3.634E+00 | 1.287E+00 | 1.207E+00 | 2.682E+00 | 9.025E+00 |
| Carbazole | 167 | 9.869E-02 | 3.564E-02 | 6.141E-02 | 2.695E-01 | 4.567E-02 | 4.869E-02 | 3.906E+00 | 8.117E+00 | 2.754E+00 | 5.438E+00 | 1.057E+01 | 7.970E-01 |
| Fluoranthene | 202 | 4.081E+00 | 2.249E+00 | 2.136E+00 | 2.961E+00 | 6.428E-01 | 6.082E-01 | 2.487E-01 | 6.368E-01 | 2.392E-01 | 3.908E-01 | 4.066E+01 | 6.647E+01 |
| Pyrene | 202 | 4.836E+00 | 3.886E+00 | 2.453E+00 | 3.276E+00 | 6.883E-01 | 5.857E-01 | 2.067E-01 | 5.428E-01 | 1.798E-01 | 3.227E-01 | 2.085E+01 | 4.113E+01 |
| C1 fluoranthene/pyrene a | 216 | 1.769E-01 | 2.398E-01 | 1.386E-01 | 1.665E-01 | 3.793E-02 | 2.421E-02 | 5.916E-01 | 1.406E+00 | 5.916E-01 | 6.421E-01 | 8.889E-01 | 3.007E+00 |
| C1 fluoranthene/pyrene b | 216 | 2.875E-01 | 3.902E-01 | 2.038E-01 | 2.838E-01 | 5.703E-02 | 3.695E-02 | 7.367E-01 | 2.078E+00 | 7.199E-01 | 8.083E-01 | 1.470E+00 | 5.731E+00 |
| C1 fluoranthene/pyrene c | 216 | 3.655E-01 | 2.707E-01 | 2.053E-01 | 2.670E-01 | 6.037E-02 | 4.840E-02 | 1.215E+00 | 3.142E+00 | 1.410E+00 | 1.690E+00 | 1.232E+00 | 7.389E+00 |
| C1 fluoranthene/pyrene d | 216 | 2.147E-01 | 2.221E-01 | 1.263E-01 | 1.923E-01 | 3.927E-02 | 2.543E-02 | 8.696E-01 | 2.275E+00 | 1.514E+00 | 1.196E+00 | 1.357E+00 | 4.105E+00 |
| C1 fluoranthene/pyrene e | 216 | 6.812E-01 | 5.112E-01 | 3.241E-01 | 3.812E-01 | 1.208E-01 | 5.716E-02 | 1.175E+00 | 3.224E+00 | 1.393E+00 | 1.783E+00 | 1.630E+00 | 5.990E+00 |
| C1 fluoranthene/pyrene f | 216 | 6.878E-01 | 6.879E-01 | 3.729E-01 | 5.278E-01 | 1.173E-01 | 7.632E-02 | 1.022E+00 | 3.049E+00 | 1.098E+00 | 1.176E+00 | 1.179E+00 | 4.696E+00 |
| Benz[a]anthracene | 228 | 5.256E-01 | 4.485E-01 | 3.228E-01 | 4.699E-01 | 8.956E-02 | 8.538E-02 | 3.779E+00 | 8.008E+00 | 3.248E+00 | 5.406E+00 | 2.399E+00 | 1.170E+01 |
| Chrysene | 228 | 6.061E-01 | 5.054E-01 | 3.750E-01 | 5.028E-01 | 8.712E-02 | 8.509E-02 | 4.593E+00 | 8.262E+00 | 3.663E+00 | 6.297E+00 | 2.951E+00 | 9.766E+00 |
| C1 Chrysene a | 242 | 5.845E-02 | 6.653E-02 | 4.143E-02 | 5.050E-02 | 1.277E-02 | 7.817E-03 | 2.653E-01 | 5.959E-01 | 2.707E-01 | 2.908E-01 | 2.649E-01 | 1.071E+00 |
| C1 Chrysene b | 242 | 1.561E-01 | 1.546E-01 | 1.852E-01 | 1.269E-01 | 3.095E-02 | 2.104E-02 | 6.722E-01 | 1.835E+00 | 7.315E-01 | 7.931E-01 | 4.108E+00 | 1.878E+00 |
| C1 Chrysene c | 242 | 1.053E-01 | 3.884E-02 | 2.849E-02 | 4.466E-02 | 1.388E-02 | 1.404E-02 | 3.212E-01 | 6.855E-01 | 7.797E-01 | 4.797E-01 | 1.967E+00 | 1.047E+00 |
| C1 Chrysene d | 242 | 9.653E-02 | 6.397E-02 | 4.186E-02 | 4.843E-02 | 7.890E-03 | 9.466E-03 | 1.598E-01 | 3.736E-01 | 2.845E-01 | 2.713E-01 | 1.189E-01 | 6.644E-01 |
| Benzo[b]fluoranthene | 252 | 1.476E-01 | 1.159E-01 | 1.082E-01 | 1.739E-01 | 2.962E-02 | 3.694E-02 | 2.118E-01 | 4.868E-01 | 1.857E-01 | 3.667E-01 | 7.895E-01 | 3.829E+00 |
| Benzo[k]fluoranthene | 252 | 2.372E-01 | 8.900E-02 | 8.750E-02 | 1.287E-01 | 3.126E-02 | 3.591E-02 | 1.487E+00 | 3.049E+00 | 1.318E+00 | 2.207E+00 | 4.833E+00 | 2.630E+00 |
| benzofluoranthene c | 252 | 3.197E-01 | 1.232E-01 | 1.098E-01 | 1.447E-01 | 4.109E-02 | 4.777E-02 | 1.431E+00 | 2.817E+00 | 1.191E+00 | 2.136E+00 | 3.962E+00 | 1.882E+00 |
| benzofluoranthene d | 252 | 1.688E-01 | 9.958E-02 | 6.471E-02 | 1.076E-01 | 2.610E-02 | 2.350E-02 | 7.220E-01 | 1.677E+00 | 6.743E-01 | 9.326E-01 | 2.187E-01 | 1.131E+00 |
| Benzo[a]pyrene | 252 | 4.803E-01 | 3.451E-01 | 2.293E-01 | 3.293E-01 | 6.712E-02 | 8.905E-02 | 3.302E+00 | 7.958E+00 | 2.386E+00 | 5.768E+00 | 7.101E+00 | 4.798E+00 |
| Benzopyrene b | 252 | 2.223E-01 | 1.674E-01 | 1.186E-01 | 1.733E-01 | 3.536E-02 | 4.285E-02 | 1.972E+00 | 4.256E+00 | 1.491E+00 | 3.170E+00 | 5.963E+00 | 2.629E+00 |
| Indeno[1,2,3-cd]pyrene | 276 | 8.906E-02 | 8.111E-02 | 5.821E-02 | 1.005E-01 | 1.803E-02 | 2.709E-02 | 1.249E+00 | 3.128E+00 | 9.398E-01 | 2.364E+00 | 2.574E+00 | 1.688E+00 |
| Dibenzo[a,h]anthracene | 278 | 2.192E-02 | 2.334E-02 | 8.333E-03 | 1.774E-02 | 2.661E-03 | 3.094E-03 | 1.825E-01 | 5.732E-01 | 1.249E-01 | 2.991E-01 | 3.303E-01 | 4.862E-01 |
| Benzo[g,h,i]perylene | 276 | 2.551E-02 | 4.007E-02 | 3.115E-02 | 6.843E-02 | 1.537E-02 | 1.822E-02 | 1.007E+00 | 1.772E+00 | 8.627E-01 | 1.906E+00 | 2.582E+00 | 1.376E+00 |
| dibenzochrysene | 276 | 2.986E-02 | 3.181E-02 | 2.124E-02 | 4.417E-02 | 4.955E-03 | 5.157E-03 | 6.393E-01 | 1.365E+00 | 2.790E-01 | 1.048E+00 | 4.822E-01 | 4.581E-01 |
| 2-phenylnaphthalene | 204 | 7.692E-02 | 4.016E-01 | 1.788E-01 | 2.554E-01 | 4.676E-02 | 2.001E-02 | 8.764E-01 | 2.076E+00 | 4.525E-01 | 9.409E-01 | 2.043E+00 | 5.269E+00 |
| phenalenothiophene a | 208 | 1.680E-01 | 3.707E-01 | 1.770E-01 | 2.628E-01 | 5.404E-02 | 3.349E-02 | 1.035E+00 | 3.064E+00 | 1.194E+00 | 1.886E+00 | 1.760E+00 | 4.828E+00 |
| phenalenothiophene b | 208 | 3.730E-02 | 2.559E-02 | 1.762E-02 | 3.975E-02 | 6.119E-03 | 2.609E-03 | 3.365E-01 | 8.854E-01 | 3.242E-01 | 5.901E-01 | 1.528E-01 | 7.876E-01 |
| anthracene, 9-ethenyl | 204 | 3.967E-02 | 5.814E-02 | 4.742E-02 | 7.616E-02 | 1.058E-02 | 3.029E-02 | 3.042E-01 | 9.466E-01 | 4.472E-01 | 5.248E-01 | 2.041E-01 | 2.241E+00 |

The header spanning columns 1–12 reads "Samples".

Table C.8: Normalised GCxGC peak areas for all coal tar samples - raw data 5/10

| Compound | Mass (g/mol) | Samples | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 2011 | 2013 | 2015 | 21 | 22 | 23 |
| undecane (C11) | 156 | 5.485E+00 | 6.311E+00 | 5.958E-03 | 1.467E-01 | 2.571E-01 | 2.503E-07 | 5.996E-02 | 2.242E-01 | 1.955E-01 | 1.489E-01 | 2.019E-01 | 4.887E-02 | 1.943E-01 |
| dodecane (C12) | 170 | 8.203E+00 | 1.050E+00 | 6.978E-03 | 1.479E-01 | 8.236E-01 | 2.503E-07 | 6.581E-02 | 3.661E-01 | 2.588E-01 | 2.516E-01 | 3.409E-01 | 6.653E-02 | 2.272E-01 |
| tridecane (C13) | 184 | 1.004E-01 | 1.502E+01 | 8.315E-03 | 1.094E-01 | 1.158E+00 | 2.503E-07 | 4.593E-02 | 3.739E-01 | 3.505E-01 | 2.559E-01 | 3.444E-01 | 5.368E-02 | 1.804E-01 |
| tetradecane (C14) | 198 | 9.143E-01 | 1.495E+01 | 1.456E-02 | 1.321E-01 | 1.243E+00 | 2.503E-07 | 2.234E-02 | 3.973E-01 | 3.470E-01 | 2.509E-01 | 3.245E-01 | 4.130E-02 | 1.478E-01 |
| pentadecane (C15) | 212 | 7.371E+00 | 1.133E-01 | 1.669E-02 | 8.657E-02 | 1.109E+00 | 2.503E-07 | 1.296E-02 | 3.871E-01 | 3.187E-01 | 2.203E-01 | 1.920E-01 | 2.773E-02 | 1.022E-01 |
| hexadecane (C16) | 226 | 5.650E+00 | 9.328E+00 | 8.982E-03 | 6.852E-02 | 9.799E-01 | 2.503E-07 | 8.191E-01 | 3.210E-01 | 2.620E-01 | 1.934E-01 | 1.769E-01 | 2.256E-02 | 7.906E-02 |
| heptadecane (C17) | 240 | 3.657E+00 | 6.618E+00 | 7.873E-03 | 3.513E-02 | 6.462E-01 | 9.691E-03 | 6.861E-01 | 2.680E-01 | 2.523E-01 | 1.773E-01 | 1.259E-01 | 1.243E-02 | 6.560E-02 |
| octadecane (C18) | 254 | 2.010E+00 | 4.387E+00 | 6.427E-03 | 2.844E-02 | 3.862E-01 | 3.254E-03 | 2.738E-03 | 2.497E-01 | 2.231E-01 | 1.636E-01 | 9.033E-02 | 7.305E-03 | 4.594E-02 |
| tetracosane (C24) | 338 | 3.563E-02 | 8.758E-02 | 8.082E-03 | 1.602E-02 | 5.164E-02 | 2.503E-07 | 2.524E-07 | 1.469E-01 | 1.174E-01 | 9.060E-02 | 2.241E-02 | 2.654E-03 | 2.004E-02 |
| pentacosane (C25) | 352 | 1.966E-02 | 5.739E-02 | 1.479E-02 | 1.181E-02 | 3.701E-02 | 2.503E-07 | 2.524E-07 | 1.357E-01 | 1.102E-01 | 8.292E-02 | 2.087E-02 | 2.706E-03 | 1.753E-02 |
| hexacosane (C26) | 366 | 1.949E-02 | 5.890E-02 | 1.722E-02 | 1.608E-02 | 3.485E-02 | 2.503E-07 | 2.524E-07 | 1.118E-01 | 9.310E-02 | 6.997E-02 | 1.627E-02 | 2.136E-03 | 1.815E-02 |
| heptacosane (C27) | 380 | 1.286E-02 | 5.143E-02 | 2.604E-02 | 1.854E-02 | 4.704E-02 | 2.503E-07 | 2.524E-07 | 9.659E-02 | 7.602E-02 | 5.420E-02 | 1.651E-02 | 2.450E-03 | 1.897E-02 |
| octacosane (C28) | 394 | 1.132E-02 | 4.239E-02 | 2.320E-02 | 9.112E-03 | 3.079E-02 | 2.503E-07 | 2.524E-07 | 7.286E-02 | 5.570E-02 | 4.206E-02 | 1.079E-02 | 1.682E-03 | 1.341E-02 |
| nonacosane (C29) | 408 | 1.016E-02 | 2.859E-02 | 1.893E-02 | 9.333E-03 | 2.267E-02 | 2.503E-07 | 2.524E-07 | 6.546E-02 | 4.928E-02 | 3.830E-02 | 7.985E-03 | 1.571E-03 | 1.097E-02 |
| triacontane (C30) | 422 | 7.722E-03 | 1.763E-02 | 1.376E-02 | 4.210E-03 | 1.676E-02 | 2.503E-07 | 2.524E-07 | 4.754E-02 | 3.059E-02 | 1.812E-02 | 6.772E-03 | 1.125E-03 | 5.241E-03 |
| hentriacontane (C31) | 436 | 2.537E-07 | 2.580E-02 | 9.402E-03 | 8.322E-04 | 2.495E-07 | 2.503E-07 | 2.524E-07 | 3.769E-02 | 2.125E-02 | 1.414E-02 | 6.055E-03 | 2.112E-03 | 3.691E-03 |
| dotriacontane (C32) | 450 | 2.537E-07 | 2.580E-07 | 8.784E-03 | 8.087E-04 | 2.495E-07 | 2.503E-07 | 2.524E-07 | 1.357E-02 | 1.110E-02 | 7.505E-03 | 3.550E-03 | 6.459E-04 | 2.128E-03 |
| tritriacontane (C33) | 464 | 2.537E-07 | 2.580E-07 | 8.305E-03 | 1.411E-03 | 2.495E-07 | 2.503E-07 | 2.524E-07 | 1.003E-02 | 8.518E-03 | 6.043E-03 | 2.983E-03 | 8.519E-04 | 1.713E-03 |
| pristane | 170 | 2.092E+00 | 2.073E+00 | 1.980E-02 | 1.005E-01 | 7.438E-01 | 4.485E-03 | 3.049E-03 | 3.304E-01 | 3.116E-01 | 2.262E-01 | 7.578E-02 | 1.849E-02 | 5.779E-02 |
| phytane | 184 | 8.690E-01 | 1.516E+00 | 3.126E-03 | 1.008E-02 | 6.129E-02 | 3.279E-03 | 2.524E-07 | 5.577E-02 | 4.874E-02 | 3.541E-02 | 1.360E-02 | 1.726E-03 | 1.271E-02 |
| branched alkane A (C12) | 198 | 4.419E+00 | 4.431E+00 | 2.619E-03 | 8.093E-02 | 7.001E-01 | 2.503E-07 | 7.930E-03 | 1.566E-01 | 1.309E-01 | 1.153E-01 | 1.689E-01 | 2.983E-02 | 7.898E-02 |
| branched alkane B (C13) | 268 | 3.647E+00 | 4.948E+00 | 9.114E-03 | 6.365E-02 | 4.077E-01 | 2.503E-07 | 1.342E-02 | 1.489E-01 | 1.367E-01 | 1.016E-01 | 8.734E-02 | 1.815E-02 | 6.403E-02 |
| branched alkane C  (C14) | 282 | 2.263E+00 | 2.665E+00 | 1.351E-02 | 6.388E-02 | 5.249E-01 | 2.503E-07 | 8.850E-03 | 1.642E-01 | 1.563E-01 | 1.147E-01 | 9.746E-02 | 1.595E-02 | 4.874E-02 |
| phenol | 94 | 2.537E-07 | 2.580E-07 | 2.469E-07 | 1.398E+00 | 2.930E-01 | 8.640E-01 | 5.794E+00 | 2.504E-07 | 2.499E-07 | 2.496E-07 | 2.305E-07 | 2.494E-07 | 3.951E-07 |
| C1 phenol a | 108 | 2.537E-07 | 6.028E-01 | 2.469E-07 | 3.055E+00 | 1.047E+00 | 3.784E-01 | 3.067E+00 | 2.504E-07 | 2.499E-07 | 2.496E-07 | 9.238E-03 | 6.080E-02 | 8.917E-02 |
| C1 phenol b | 108 | 2.537E-07 | 2.625E-02 | 2.469E-07 | 1.473E+00 | 2.051E+00 | 3.129E-01 | 3.740E+00 | 2.504E-07 | 2.499E-07 | 2.496E-07 | 1.270E-03 | 7.613E-03 | 6.416E-02 |
| C2 phenol a | 122 | 2.537E-07 | 3.743E-01 | 2.469E-07 | 5.900E-01 | 2.932E-01 | 5.264E-02 | 3.510E-01 | 4.030E-03 | 3.931E-03 | 4.545E-03 | 2.305E-07 | 6.309E-02 | 4.593E-02 |
| C2 phenol b | 122 | 2.537E-07 | 8.185E-01 | 2.469E-07 | 1.854E+00 | 9.177E-01 | 1.872E-01 | 1.554E+00 | 1.749E-02 | 1.924E-02 | 1.956E-02 | 2.305E-07 | 1.853E-01 | 2.016E-01 |
| C2 phenol c | 122 | 2.537E-07 | 6.357E-02 | 2.469E-07 | 2.384E-01 | 1.016E-01 | 4.118E-02 | 4.741E-01 | 8.127E-03 | 1.020E-02 | 1.409E-02 | 2.305E-07 | 9.797E-02 | 8.096E-02 |
| C2 phenol d | 122 | 2.537E-07 | 9.760E-02 | 2.469E-07 | 2.252E-01 | 1.236E-01 | 1.722E-02 | 1.151E-01 | 5.685E-03 | 6.505E-03 | 4.800E-03 | 2.305E-07 | 8.752E-02 | 3.545E-02 |
| toluene | 91 | 5.368E-02 | 8.198E+00 | 8.382E+00 | 8.614E-01 | 1.451E+00 | 9.800E-01 | 1.545E+00 | 5.770E-01 | 1.285E-01 | 1.035E-01 | 9.062E+00 | 2.001E+00 | 6.736E-01 |
| C3-enyl benzene | 120 | 5.916E+01 | 3.964E+01 | 4.650E-02 | 1.128E+00 | 9.142E-01 | 7.025E-01 | 3.160E+00 | 9.642E-02 | 8.353E-01 | 1.353E-01 | 6.233E-01 | 9.611E-01 | 4.816E-01 |
| C3 benzene a | 120 | 8.843E+01 | 7.216E+01 | 6.373E-02 | 4.388E+00 | 4.537E+00 | 7.773E-01 | 5.313E+00 | 2.049E-01 | 2.086E-01 | 1.242E-01 | 3.165E+00 | 2.200E+00 | 2.427E-01 |
| C3 benzene b | 120 | 3.196E+01 | 2.492E+01 | 3.014E-02 | 1.544E+00 | 1.930E+00 | 1.736E-01 | 1.362E+00 | 2.465E-01 | 2.388E-01 | 1.712E-01 | 7.659E-01 | 2.986E-01 | 7.374E-01 |
| C4 benzene a | 134 | 3.906E+00 | 3.902E+01 | 4.740E-02 | 1.223E+00 | 8.962E-01 | 7.025E-01 | 3.160E+00 | 3.079E-01 | 2.396E-01 | 1.321E-01 | 6.498E-02 | 2.116E-03 | 8.821E-03 |
| C4 benzene b | 134 | 9.775E+00 | 5.383E+00 | 4.607E-03 | 4.186E-01 | 9.503E-01 | 2.141E-02 | 3.207E-02 | 2.058E-01 | 1.679E-01 | 1.329E-01 | 3.848E-01 | 3.769E-02 | 5.268E-02 |
| C4 benzene c | 134 | 1.233E-01 | 7.219E+00 | 4.499E-03 | 3.740E-01 | 1.007E+00 | 1.236E-02 | 6.491E-02 | 4.126E-01 | 2.578E-01 | 1.680E-01 | 3.270E-01 | 1.918E-02 | 1.345E-02 |
| C4 benzene d | 134 | 1.380E-01 | 1.031E+01 | 8.033E-03 | 7.016E-01 | 1.231E+00 | 1.252E-02 | 3.578E-01 | 6.743E-02 | 4.774E-02 | 7.208E-02 | 2.381E-01 | 1.389E-02 | 8.357E-02 |
| C4 benzene e | 134 | 1.148E+01 | 9.897E+00 | 2.469E-07 | 5.725E-01 | 1.109E+00 | 8.399E-02 | 5.136E-01 | 5.513E-02 | 5.428E-01 | 3.797E-02 | 3.441E-01 | 3.361E-02 | 4.783E-02 |
| C4 benzene f | 134 | 1.905E+01 | 1.585E+01 | 1.507E-02 | 4.874E-01 | 1.454E+00 | 1.062E-01 | 8.198E-01 | 6.933E-02 | 6.959E-02 | 4.969E-02 | 5.537E-01 | 5.061E-02 | 5.780E-02 |
| C4 benzene g | 134 | 2.032E-01 | 1.664E+01 | 1.080E-02 | 7.250E-01 | 1.675E+00 | 4.406E-02 | 3.405E-01 | 4.238E-02 | 2.870E-02 | 3.761E-02 | 9.098E-02 | 3.140E-02 | 5.496E-02 |

272

| Compound | Mass (g/mol) | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 2011 | 2013 | 2015 | 21 | 22 | 23 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| indane | 118 | 3.530E-01 | 4.826E+00 | 1.163E-02 | 8.272E-01 | 1.450E+00 | 3.018E-01 | 1.329E+00 | 1.873E-01 | 1.806E-01 | 1.420E-01 | 2.127E-01 | 9.822E-02 | 1.150E-01 |
| indene | 116 | 8.713E-01 | 1.145E+01 | 1.977E-01 | 1.942E-01 | 3.664E+01 | 3.759E+01 | 8.640E+01 | 1.012E+01 | 8.866E-01 | 7.866E-01 | 1.302E+00 | 3.740E+00 | 1.252E+00 |
| benzofuran | 118 | 2.153E-01 | 1.041E+01 | 2.966E-02 | 6.974E+00 | 2.433E+00 | 1.459E+00 | 7.027E+00 | 1.725E-01 | 1.549E-01 | 1.267E-01 | 4.381E-01 | 9.611E-01 | 9.711E-02 |
| benzothiophene | 134 | 1.708E-02 | 1.874E+02 | 9.899E-02 | 1.295E-01 | 5.134E+01 | 5.556E+01 | 8.934E+01 | 1.765E+00 | 1.543E+00 | 1.427E+00 | 2.107E+00 | 2.632E+00 | 8.125E-01 |
| C1 benzothiophene a | 148 | 6.193E-01 | 8.129E+01 | 1.196E-02 | 1.835E-02 | 8.020E+00 | 1.583E+00 | 4.248E+00 | 3.340E-01 | 3.176E-01 | 2.771E-01 | 5.203E-01 | 1.654E-01 | 1.914E-01 |
| C1 benzothiophene b | 148 | 2.162E-01 | 3.442E+01 | 1.780E-03 | 1.332E-02 | 4.407E+00 | 1.744E+00 | 4.307E+00 | 1.684E-01 | 1.753E-01 | 1.370E-01 | 3.149E-01 | 2.112E-01 | 6.203E-02 |
| C1 benzothiophene c | 148 | 5.343E-01 | 6.957E+01 | 1.690E-02 | 1.217E-02 | 7.094E+00 | 8.096E-01 | 2.858E+00 | 2.085E-01 | 2.463E-01 | 1.791E-01 | 4.539E-01 | 1.371E-01 | 1.537E-01 |
| C2 benzothiophene a | 162 | 1.707E-01 | 1.973E+01 | 3.637E-03 | 1.944E-01 | 1.285E+00 | 7.314E-02 | 2.567E-01 | 6.615E-02 | 6.806E-02 | 4.939E-02 | 1.416E-01 | 2.781E-02 | 3.850E-02 |
| C2 benzothiophene b | 162 | 5.783E+00 | 6.426E+01 | 2.549E-03 | 1.569E-01 | 7.173E+00 | 7.284E-02 | 3.341E-01 | 6.298E-02 | 5.256E-02 | 3.516E-02 | 1.345E-01 | 3.258E-02 | 2.767E-03 |
| C2 benzothiophene c | 162 | 9.927E+00 | 1.300E+01 | 2.678E-04 | 2.195E-01 | 9.668E-01 | 4.495E-02 | 2.848E-01 | 5.201E-02 | 9.102E-02 | 6.712E-02 | 2.607E-01 | 8.349E-02 | 4.736E-02 |
| C2 benzothiophene d | 162 | 2.564E-01 | 2.803E+01 | 9.633E-03 | 2.370E-01 | 1.601E+00 | 1.110E-01 | 3.673E-01 | 1.289E-01 | 1.038E-01 | 8.869E-02 | 2.700E-01 | 2.179E-02 | 7.895E-02 |
| C2 benzothiophene e | 162 | 1.828E-01 | 2.112E+01 | 8.766E-03 | 1.872E-01 | 1.211E+00 | 4.222E-02 | 1.576E-01 | 3.266E-02 | 5.778E-02 | 4.826E-02 | 1.787E-01 | 3.917E-02 | 3.683E-02 |
| C2 benzothiophene f | 162 | 1.384E-01 | 1.753E+01 | 5.304E-03 | 1.777E-01 | 1.009E+00 | 8.567E-02 | 3.469E-01 | 6.711E-02 | 6.792E-02 | 4.917E-02 | 7.260E-02 | 1.443E-02 | 2.393E-02 |
| C2 benzothiophene g | 162 | 1.376E-01 | 1.311E+01 | 3.055E-03 | 9.564E-01 | 5.628E-01 | 2.489E-02 | 7.576E-02 | 3.460E-02 | 2.771E-02 | 2.045E-02 | 2.590E-02 | 5.237E-03 | 4.950E-03 |
| C2 benzothiophene h | 162 | 1.273E+00 | 2.015E+00 | 8.642E-04 | 4.119E-02 | 2.298E-01 | 1.378E-02 | 6.112E-02 | 3.100E-02 | 2.232E-02 | 3.315E-02 | 2.809E-02 | 4.965E-03 | 2.842E-03 |
| C2 benzothiophene i | 162 | 8.174E-01 | 1.490E+00 | 4.204E-04 | 1.621E-02 | 8.067E-02 | 3.706E-03 | 2.100E-02 | 1.841E-02 | 1.225E-02 | 8.197E-03 | 1.167E-02 | 2.109E-03 | 4.839E-03 |
| C3 benzothiophene a | 176 | 1.668E+00 | 1.672E+00 | 2.392E-03 | 3.968E-02 | 1.690E-01 | 1.416E-02 | 5.150E-02 | 2.066E-02 | 2.013E-02 | 1.287E-02 | 7.120E-02 | 1.171E-02 | 4.705E-03 |
| C3 benzothiophene b | 176 | 1.580E-01 | 1.068E+01 | 4.416E-03 | 4.421E-02 | 7.389E-01 | 1.499E-02 | 2.660E-02 | 5.691E-02 | 3.746E-02 | 3.847E-02 | 7.685E-02 | 9.382E-03 | 2.114E-02 |
| C3 benzothiophene c | 176 | 1.965E+00 | 2.174E+00 | 1.859E-03 | 3.289E-02 | 8.170E-02 | 1.599E-02 | 3.255E-02 | 4.718E-02 | 4.717E-02 | 3.893E-02 | 1.209E-01 | 8.544E-03 | 2.956E-02 |
| C3 benzothiophene d | 176 | 1.576E-01 | 1.205E+01 | 3.703E-03 | 7.483E-02 | 7.437E-01 | 9.153E-03 | 5.157E-02 | 3.435E-02 | 3.317E-02 | 2.422E-02 | 7.570E-02 | 1.366E-02 | 4.080E-02 |
| C3 benzothiophene e | 176 | 6.847E+00 | 4.286E+00 | 1.992E-03 | 3.643E-02 | 3.785E-01 | 9.366E-03 | 2.213E-02 | 2.391E-02 | 2.136E-02 | 1.185E-02 | 2.762E-02 | 1.127E-02 | 1.949E-02 |
| diphenylmethane | 167 | 4.955E+00 | 2.494E+00 | 2.469E-07 | 1.185E-01 | 3.587E-01 | 1.846E-01 | 5.409E-01 | 3.101E-02 | 2.906E-02 | 2.164E-02 | 1.842E-02 | 3.280E-02 | 8.950E-03 |
| 1,2-dihydronaphthalene | 130 | 4.442E-01 | 4.090E+01 | 1.762E-02 | 2.556E+00 | 3.698E+00 | 8.510E-01 | 3.486E+00 | 2.328E-01 | 2.203E-01 | 1.682E-01 | 2.766E-01 | 4.506E-01 | 2.783E-01 |
| 1,4-dihydronaphthalene | 130 | 4.645E-01 | 4.122E+01 | 5.336E-02 | 3.163E+00 | 4.696E+00 | 1.335E+00 | 4.657E+00 | 2.187E-01 | 1.960E-01 | 1.522E-01 | 8.440E-01 | 6.529E-01 | 3.949E-01 |
| tetrahydronaphthalene | 132 | 8.572E-01 | 5.860E-01 | 2.469E-07 | 2.798E-02 | 8.542E-02 | 1.070E-02 | 3.080E-02 | 1.022E-02 | 9.071E-03 | 6.662E-03 | 1.084E-02 | 4.376E-03 | 8.057E-03 |
| naphthalene | 128 | 3.351E+02 | 2.384E+02 | 2.192E+00 | 7.360E+01 | 1.193E+02 | 2.443E+02 | 3.327E+02 | 2.506E+00 | 4.219E+00 | 3.742E+00 | 2.291E+01 | 1.423E+01 | 3.940E+00 |
| 2-methyl naphthalene | 142 | 2.712E+02 | 2.406E+02 | 1.358E-01 | 3.436E-01 | 9.093E+01 | 7.296E+01 | 1.425E+01 | 1.833E+00 | 1.677E+00 | 1.574E+00 | 1.169E+01 | 5.944E+00 | 1.441E+00 |
| 1-methyl naphthalene | 142 | 1.729E-02 | 1.852E+02 | 4.927E-01 | 2.042E-01 | 5.499E+01 | 3.255E+01 | 7.291E+01 | 1.197E+00 | 1.067E+00 | 9.112E-01 | 8.615E+00 | 3.822E+00 | 9.954E-01 |
| 2-EtN | 156 | 2.731E-01 | 2.012E+01 | 1.480E-02 | 9.829E-01 | 3.098E+00 | 2.763E-01 | 1.238E+00 | 3.687E-01 | 3.687E-01 | 4.002E-01 | 1.733E+00 | 3.068E-01 | 1.598E-01 |
| 2,6-DMN & 2,7-DMN | 156 | 1.066E-02 | 8.885E+01 | 1.127E-01 | 5.137E-01 | 1.227E+00 | 4.516E-01 | 1.564E+00 | 8.050E-01 | 6.702E-01 | 6.086E-01 | 5.101E+00 | 1.410E-01 | 4.011E-01 |
| 1-EtN | 156 | 1.829E-01 | 1.076E+01 | 2.507E-02 | 4.066E-02 | 1.477E+00 | 2.861E-01 | 5.613E-01 | 1.759E-01 | 1.893E-01 | 1.454E-01 | 5.179E-01 | 2.276E-01 | 6.108E-02 |
| 1,6-DMN | 156 | 4.111E+02 | 4.491E+01 | 1.311E-01 | 5.888E+00 | 1.498E+01 | 3.224E+00 | 1.237E+00 | 5.497E-01 | 7.227E-01 | 5.699E-01 | 4.773E+00 | 1.142E+00 | 3.514E-01 |
| 1,7-DMN & 1,3-DMN | 156 | 1.054E+02 | 8.404E+01 | 2.266E-01 | 4.029E-01 | 9.643E+00 | 2.594E+00 | 8.365E+00 | 8.831E-01 | 6.270E-01 | 4.947E-01 | 4.899E+00 | 1.380E+00 | 3.691E-01 |
| 2,3-DMN | 156 | 2.791E-01 | 2.378E+01 | 3.709E-02 | 1.644E-01 | 4.105E+00 | 8.651E-01 | 3.358E+00 | 4.234E-01 | 4.168E-01 | 4.013E-01 | 1.478E+00 | 4.484E-01 | 1.124E-01 |
| 1,4-DMN | 156 | 2.903E-01 | 2.149E+01 | 4.626E-02 | 1.030E-02 | 2.714E+00 | 4.319E-01 | 1.792E+00 | 1.980E-01 | 1.741E-01 | 1.522E-01 | 8.259E-01 | 2.513E-01 | 1.154E-01 |
| 1,5-DMN | 156 | 2.052E+01 | 1.664E+01 | 4.465E-02 | 9.236E-01 | 2.174E+00 | 4.165E-01 | 9.734E+00 | 2.095E-01 | 2.055E-01 | 1.551E-01 | 9.393E-01 | 1.935E-01 | 7.554E-02 |
| 1,2-DMN | 156 | 3.557E-01 | 2.898E+01 | 5.148E-02 | 2.019E-01 | 4.071E+00 | 4.311E-01 | 2.127E+00 | 3.112E-01 | 2.312E-01 | 2.127E-01 | 1.313E+00 | 2.875E-01 | 1.084E-01 |
| 1,8-DMN | 156 | 1.650E+00 | 1.459E+00 | 5.803E-04 | 4.099E-02 | 8.045E-02 | 2.934E-02 | 3.349E-02 | 6.140E-03 | 5.084E-03 | 4.440E-03 | 2.808E-02 | 1.653E-02 | 2.809E-03 |
| C3 alkyl naphthalene a | 170 | 6.273E+00 | 4.024E+00 | 4.430E-04 | 9.934E-02 | 2.070E-01 | 2.508E-07 | 1.201E-02 | 1.007E-01 | 9.580E-02 | 7.438E-02 | 8.277E-02 | 1.723E-02 | 3.869E-02 |
| C3 alkyl naphthalene b | 170 | 7.624E+00 | 4.291E+00 | 4.405E-02 | 6.806E-01 | 1.742E+00 | 2.185E-02 | 1.600E-02 | 3.204E-01 | 2.956E-01 | 2.122E-01 | 7.644E-01 | 1.652E-01 | 1.217E-01 |
| C3 alkyl naphthalene c | 170 | 3.374E+00 | 1.761E+00 | 2.685E-02 | 1.057E-01 | 6.669E-01 | 4.111E-02 | 1.433E-01 | 2.423E-01 | 1.745E-01 | 1.278E-01 | 1.177E+00 | 3.851E-02 | 7.817E-02 |

273

Table C.10: Normalised GCxGC peak areas for all coal tar samples - raw data 7/10

| Compound | Mass (g/mol) | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 2011 | 2013 | 2015 | 21 | 22 | 23 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| C3 alkyl naphthalene d | 170 | 4.634E-01 | 2.822E+01 | 9.860E-02 | 2.437E+00 | 4.251E+00 | 7.739E-01 | 3.056E-01 | 5.579E-01 | 6.258E-01 | 4.507E-01 | 1.358E+00 | 5.604E-01 | 8.028E-02 |
| C3 alkyl naphthalene e | 170 | 3.269E+01 | 2.132E+01 | 1.169E-01 | 1.831E+00 | 4.422E+00 | 5.737E-01 | 2.341E+00 | 4.556E-01 | 5.072E-01 | 4.079E-01 | 2.220E+00 | 4.734E-01 | 1.289E-01 |
| C3 alkyl naphthalene f | 170 | 1.468E-01 | 1.021E+01 | 6.078E-02 | 1.029E+00 | 1.823E+00 | 2.032E-01 | 1.040E+00 | 2.594E-01 | 2.710E-01 | 2.061E-01 | 1.062E+00 | 1.972E-01 | 7.097E-02 |
| C3 alkyl naphthalene g | 170 | 6.703E+00 | 4.047E+00 | 4.435E-02 | 8.951E-01 | 1.282E+00 | 6.995E-02 | 2.127E-01 | 1.116E-01 | 1.353E-01 | 1.005E-01 | 4.230E-01 | 1.084E-01 | 2.190E-02 |
| C3 alkyl naphthalene h | 170 | 3.145E+00 | 3.553E+00 | 8.920E-03 | 2.000E-01 | 3.998E-01 | 2.428E-02 | 6.988E-02 | 2.367E-01 | 1.783E-01 | 1.398E-01 | 9.005E-01 | 1.097E-01 | 4.470E-02 |
| C3 alkyl naphthalene i | 170 | 6.292E+00 | 5.209E+00 | 5.164E-02 | 8.000E-01 | 9.370E-01 | 4.999E-02 | 2.512E-01 | 5.693E-02 | 5.039E-02 | 3.996E-02 | 1.607E-01 | 3.405E-02 | 2.159E-02 |
| biphenyl | 154 | 1.080E-02 | 8.855E+01 | 1.233E-01 | 6.154E-01 | 2.264E+01 | 2.762E-01 | 5.329E-01 | 1.199E+00 | 1.157E+00 | 9.276E-01 | 2.218E+00 | 1.780E+00 | 3.401E-01 |
| C1 biphenyl a | 168 | 2.666E-01 | 1.770E+01 | 2.613E-02 | 1.255E+00 | 3.787E+00 | 3.186E+00 | 8.416E+00 | 3.634E-01 | 3.693E-01 | 2.628E-01 | 9.024E-01 | 4.677E-01 | 8.787E-02 |
| C1 biphenyl b | 168 | 1.477E-01 | 1.001E+01 | 1.408E-02 | 6.456E-01 | 1.782E+00 | 2.107E+00 | 4.584E+00 | 1.225E-01 | 1.586E-01 | 1.634E-01 | 4.862E-01 | 2.697E-01 | 5.011E-02 |
| dibenzofuran | 168 | 1.564E-02 | 4.770E+01 | 4.970E-01 | 2.049E+01 | 6.699E+01 | 1.603E+02 | 2.271E+01 | 3.351E+00 | 2.989E+00 | 2.869E+00 | 2.031E+01 | 6.452E+00 | 2.986E+00 |
| C1 dibenzofuran a | 182 | 1.828E-01 | 8.005E+00 | 1.309E-01 | 2.463E+00 | 6.078E+00 | 8.373E+00 | 1.913E+01 | 6.779E-01 | 6.122E-01 | 5.201E-01 | 4.680E+00 | 1.043E+00 | 6.632E-02 |
| C1 dibenzofuran b | 182 | 2.089E-01 | 8.017E+00 | 1.428E-01 | 4.240E+00 | 9.002E+00 | 1.138E+01 | 2.671E+01 | 7.747E-01 | 9.810E-01 | 7.737E-01 | 7.509E+00 | 1.662E+00 | 8.020E-02 |
| C1 dibenzofuran c | 182 | 4.633E+00 | 2.520E+00 | 3.308E-02 | 1.058E+00 | 1.862E+00 | 1.458E+00 | 4.187E+00 | 2.147E-01 | 1.873E-01 | 1.171E-01 | 1.378E+00 | 3.449E-01 | 2.271E-02 |
| acenaphthylene | 152 | 1.774E-02 | 2.575E+02 | 7.483E-01 | 3.169E-01 | 9.873E-01 | 2.236E+02 | 1.858E+02 | 1.785E+00 | 1.616E+00 | 1.485E+00 | 8.679E+00 | 5.795E+00 | 1.319E+00 |
| acenaphthene | 154 | 6.699E-01 | 2.732E-01 | 1.013E-01 | 1.943E-01 | 1.356E-01 | 5.845E+00 | 1.297E-01 | 1.055E+00 | 1.198E+00 | 1.080E+00 | 6.580E+00 | 8.904E-01 | 1.225E-01 |
| dibenzothiophene | 184 | 9.399E-01 | 1.660E+02 | 1.113E-01 | 1.870E+00 | 1.583E+01 | 3.184E+01 | 5.092E+01 | 1.161E+00 | 1.032E+00 | 9.399E-01 | 4.996E+00 | 7.734E-01 | 2.647E-01 |
| naphtho-thiophene | 184 | 7.975E+00 | 9.435E+00 | 2.801E-02 | 6.433E-01 | 6.868E+00 | 5.854E+00 | 9.107E+00 | 8.942E-01 | 8.113E-01 | 6.365E-01 | 1.680E+00 | 3.949E-01 | 7.758E-02 |
| C1 dibenzothiophene a | 198 | 3.788E-01 | 5.358E+01 | 2.075E-02 | 1.281E-01 | 2.314E+00 | 1.087E+00 | 2.558E+00 | 7.137E-02 | 6.487E-02 | 5.348E-02 | 5.390E-01 | 1.010E-01 | 9.529E-02 |
| C1 dibenzothiophene b | 198 | 1.594E-01 | 2.366E+01 | 1.025E-02 | 1.582E-01 | 1.192E+00 | 7.141E-01 | 1.756E+00 | 7.700E-02 | 7.327E-02 | 5.411E-02 | 8.804E-01 | 1.735E-01 | 9.215E-02 |
| C1 dibenzothiophene c | 198 | 1.399E-01 | 2.374E+01 | 1.140E-02 | 1.859E-01 | 1.592E+00 | 8.013E-01 | 1.615E+00 | 4.100E-02 | 5.011E-02 | 5.648E-02 | 6.328E-01 | 8.982E-02 | 5.058E-02 |
| C2 dibenzothiophene a | 212 | 7.358E-01 | 7.080E-01 | 5.305E-04 | 7.265E-03 | 1.047E-01 | 8.438E-02 | 1.945E-01 | 6.741E-03 | 5.724E-03 | 2.112E-03 | 7.219E-02 | 1.318E-02 | 2.135E-02 |
| C2 dibenzothiophene b | 212 | 8.193E-01 | 1.082E-01 | 3.939E-03 | 1.046E-02 | 2.835E-01 | 4.967E-02 | 1.225E-01 | 1.101E-02 | 5.970E-03 | 5.459E-03 | 7.823E-02 | 1.385E-02 | 1.172E-02 |
| C2 dibenzothiophene c | 212 | 3.354E+00 | 4.450E+00 | 4.776E-03 | 1.759E-02 | 1.866E+00 | 1.325E-01 | 2.739E-01 | 1.395E-02 | 1.391E-02 | 1.084E-02 | 1.082E-01 | 1.958E-02 | 1.284E-02 |
| C2 dibenzothiophene d | 212 | 4.502E+00 | 6.009E+00 | 4.217E-03 | 2.046E-02 | 2.358E-01 | 1.227E-01 | 5.233E-01 | 1.894E-02 | 2.390E-02 | 2.061E-02 | 1.105E-01 | 2.076E-02 | 1.856E-02 |
| C2 dibenzothiophene e | 212 | 6.502E+00 | 9.090E+00 | 3.890E-03 | 2.069E-02 | 3.727E-01 | 9.366E-02 | 2.018E-01 | 8.678E-03 | 8.373E-03 | 2.683E-02 | 1.513E-01 | 2.682E-02 | 1.918E-02 |
| C2 dibenzothiophene f | 212 | 5.594E-01 | 1.100E+00 | 1.247E-03 | 1.037E-02 | 4.696E-02 | 1.194E-02 | 2.229E-02 | 2.226E-02 | 4.329E-02 | 2.908E-02 | 1.545E-01 | 3.894E-02 | 1.699E-02 |
| C3 dibenzothiophene a | 226 | 1.805E+00 | 2.022E+00 | 2.469E-07 | 5.918E-03 | 5.671E-02 | 1.698E-02 | 5.347E-02 | 1.352E-03 | 1.387E-03 | 6.800E-04 | 3.185E-03 | 4.974E-03 | 9.857E-04 |
| C3 dibenzothiophene b | 226 | 6.584E-01 | 1.012E+00 | 2.469E-07 | 4.076E-03 | 4.024E-02 | 2.191E-02 | 4.599E-02 | 3.613E-03 | 3.009E-03 | 3.989E-03 | 2.491E-02 | 8.359E-03 | 6.509E-03 |
| C3 dibenzothiophene c | 226 | 5.046E-01 | 5.986E-01 | 2.469E-07 | 2.407E-03 | 2.993E-02 | 1.088E-02 | 3.069E-02 | 4.820E-03 | 3.475E-03 | 3.223E-02 | 3.106E-02 | 2.150E-03 | 5.044E-03 |
| C3 dibenzothiophene d | 226 | 5.887E-01 | 8.157E-01 | 2.469E-07 | 2.456E-03 | 2.368E-02 | 7.276E-03 | 1.037E-02 | 4.135E-03 | 1.056E-02 | 3.440E-02 | 2.043E-02 | 2.405E-03 | 1.757E-03 |
| C3 dibenzothiophene e | 226 | 9.870E-01 | 1.120E+00 | 7.466E-04 | 1.855E-02 | 8.384E-02 | 9.334E-03 | 2.208E-02 | 3.953E-03 | 4.915E-03 | 5.309E-03 | 8.945E-03 | 1.992E-03 | 2.308E-03 |
| C3 dibenzothiophene f | 226 | 1.769E+00 | 2.385E+00 | 6.510E-04 | 5.297E-03 | 8.843E-02 | 1.945E-02 | 1.056E-01 | 1.471E-02 | 7.749E-03 | 6.226E-03 | 1.427E-02 | 2.516E-03 | 1.028E-03 |
| C3 dibenzothiophene g | 226 | 9.194E-01 | 1.003E+00 | 1.052E-03 | 1.310E-02 | 5.543E-02 | 7.965E-03 | 1.319E-02 | 4.817E-03 | 5.777E-03 | 5.018E-03 | 1.798E-02 | 2.080E-03 | 1.521E-03 |
| C3 dibenzothiophene h | 226 | 3.080E-01 | 4.874E-01 | 5.093E-04 | 6.257E-03 | 2.737E-02 | 9.417E-03 | 1.159E-02 | 5.077E-03 | 5.329E-03 | 4.187E-03 | 1.489E-02 | 2.317E-03 | 9.541E-04 |
| fluorene | 166 | 1.234E-02 | 1.326E+02 | 5.437E-01 | 1.453E-01 | 5.142E-01 | 8.512E-01 | 1.602E-01 | 1.963E+00 | 1.781E+00 | 1.670E+00 | 1.106E+01 | 4.336E+00 | 5.095E-01 |
| C1 fluorene a | 180 | 6.400E+00 | 5.510E+00 | 5.071E-03 | 5.050E-01 | 9.546E-01 | 8.520E-01 | 1.465E+00 | 2.829E-01 | 3.161E-01 | 2.789E-01 | 3.066E+00 | 7.177E-01 | 8.963E-02 |
| C1 fluorene b | 180 | 2.444E+00 | 2.373E+00 | 4.475E-03 | 2.231E-01 | 2.820E-01 | 3.488E-01 | 8.536E-01 | 1.602E-01 | 1.275E-01 | 9.893E-01 | 1.028E+00 | 4.239E-01 | 6.545E-02 |
| phenanthrene | 178 | 2.697E+02 | 2.874E+02 | 2.513E+00 | 4.767E-01 | 1.927E+02 | 4.093E+02 | 5.321E+02 | 7.937E+00 | 6.759E+00 | 6.706E+00 | 3.577E+01 | 1.056E+01 | 1.545E+00 |
| anthracene | 178 | 7.879E+01 | 7.849E+01 | 8.543E-01 | 1.643E-01 | 5.885E+01 | 1.579E+02 | 1.991E+02 | 3.191E+00 | 3.332E+00 | 5.730E+00 | 1.764E+01 | 5.689E+00 | 5.067E-01 |
| C1 phenanthrene a | 192 | 2.599E-01 | 2.611E+01 | 1.427E-01 | 2.088E+00 | 5.287E+00 | 6.966E+00 | 1.663E+01 | 5.717E-01 | 5.522E-01 | 4.664E-01 | 4.670E+00 | 1.143E+00 | 1.437E-01 |
| C1 phenanthrene b | 192 | 2.904E-01 | 3.044E+01 | 1.391E-01 | 2.950E+00 | 6.441E+00 | 9.270E+00 | 2.199E+01 | 8.262E-01 | 6.811E-01 | 6.392E-01 | 7.509E+00 | 1.508E+00 | 1.618E-01 |

| Compound | Mass (g/mol) | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 2011 | 2013 | 2015 | 21 | 22 | 23 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| C1 phenanthrene c | 192 | 1.273E-01 | 1.143E+01 | 1.066E-01 | 2.261E+00 | 3.486E+00 | 3.630E+00 | 7.904E+00 | 3.336E-01 | 3.714E-01 | 3.996E-01 | 1.932E+00 | 6.339E-01 | 7.368E-02 |
| C1 anthracene a | 192 | 7.924E+00 | 6.138E+00 | 7.125E-02 | 1.383E+00 | 1.657E+00 | 1.774E+00 | 2.940E+00 | 2.820E-01 | 2.411E-01 | 1.797E-01 | 1.364E+00 | 3.144E-01 | 5.059E-02 |
| C1 anthracene b | 192 | 2.334E-01 | 1.634E+01 | 1.286E-01 | 1.113E+00 | 2.669E+00 | 2.617E+00 | 8.055E+00 | 2.861E-01 | 2.985E-01 | 2.255E-01 | 2.393E+00 | 4.525E-01 | 8.782E-02 |
| C1 anthracene c | 192 | 1.934E+01 | 1.790E+01 | 1.170E-01 | 1.325E+00 | 3.619E+00 | 3.424E+00 | 7.605E+00 | 3.765E-01 | 3.511E-01 | 2.832E-01 | 2.804E+00 | 7.272E-01 | 1.246E-01 |
| Carbazole | 167 | 3.770E+00 | 5.880E+00 | 2.649E-02 | 1.474E+00 | 9.149E+00 | 4.576E+01 | 8.492E+01 | 1.293E+00 | 1.249E+00 | 1.461E+00 | 1.623E+00 | 2.395E+00 | 5.903E-02 |
| Fluoranthene | 202 | 5.647E+01 | 8.898E+01 | 1.393E+00 | 1.842E+01 | 9.964E+01 | 2.218E+02 | 3.181E+02 | 4.143E+00 | 3.472E+00 | 2.731E+00 | 2.355E+01 | 2.887E+00 | 5.251E-01 |
| Pyrene | 202 | 6.909E+01 | 1.031E+02 | 1.388E+00 | 1.508E+01 | 8.383E+01 | 1.770E+02 | 2.165E+02 | 3.772E+00 | 3.365E+00 | 2.625E+00 | 2.047E+01 | 4.306E+00 | 6.615E-01 |
| C1 fluoranthene/pyrene a | 216 | 4.820E+00 | 5.102E+00 | 6.839E-02 | 8.476E-01 | 2.020E+00 | 3.332E+00 | 7.406E+00 | 2.726E-01 | 2.481E-01 | 1.869E-01 | 2.037E+00 | 3.769E-01 | 4.082E-02 |
| C1 fluoranthene/pyrene b | 216 | 7.471E+00 | 7.813E+00 | 8.623E-02 | 1.130E+00 | 2.223E+00 | 4.297E+00 | 8.393E+00 | 2.600E-01 | 2.175E-01 | 1.641E-01 | 1.705E+00 | 8.962E-01 | 5.342E-02 |
| C1 fluoranthene/pyrene c | 216 | 6.184E+00 | 7.534E+00 | 1.407E-01 | 1.825E+00 | 4.828E+00 | 1.482E+01 | 3.001E+01 | 5.419E-01 | 5.417E-01 | 4.355E-01 | 5.182E+00 | 1.637E+00 | 8.478E-02 |
| C1 fluoranthene/pyrene d | 216 | 3.727E+00 | 6.321E+00 | 1.350E-01 | 9.521E-01 | 6.097E+00 | 1.247E+00 | 2.376E+01 | 3.089E-01 | 2.937E-01 | 2.811E-01 | 2.837E+00 | 8.373E-01 | 5.290E-02 |
| C1 fluoranthene/pyrene e | 216 | 8.836E+00 | 1.002E+01 | 1.031E-01 | 1.799E+00 | 1.972E+00 | 9.280E+00 | 1.981E+01 | 4.011E-01 | 3.233E-01 | 2.951E-01 | 3.939E+00 | 6.685E-01 | 9.023E-02 |
| C1 fluoranthene/pyrene f | 216 | 1.072E+01 | 1.074E+01 | 2.163E-01 | 1.597E+00 | 3.559E+00 | 1.204E+01 | 1.484E+01 | 4.367E-01 | 3.904E-01 | 2.948E-01 | 1.720E+00 | 3.619E-01 | 1.267E-01 |
| Benz[a]anthracene | 228 | 7.552E+00 | 1.344E+01 | 2.146E-01 | 3.442E+00 | 1.253E+01 | 4.576E+01 | 6.693E+01 | 1.443E+00 | 1.292E+00 | 9.930E-01 | 8.357E+00 | 2.017E+00 | 1.586E-01 |
| Chrysene | 228 | 7.735E-01 | 1.414E+01 | 2.661E-01 | 3.214E+00 | 1.401E+01 | 4.743E+01 | 6.648E+01 | 1.408E+00 | 1.157E+00 | 9.805E-01 | 7.561E+00 | 1.802E+00 | 1.728E-01 |
| C1 Chrysene a | 242 | 3.080E-01 | 3.711E-01 | 2.519E-02 | 3.430E-01 | 1.060E+00 | 1.401E+00 | 6.047E+00 | 1.149E+00 | 8.436E-03 | 5.043E-03 | 6.629E-02 | 2.422E-02 | 9.049E-03 |
| C1 Chrysene b | 242 | 8.227E-01 | 1.176E+00 | 2.636E-02 | 3.414E-01 | 5.520E-01 | 1.010E+00 | 2.252E+00 | 1.456E-02 | 1.107E-02 | 7.897E-03 | 1.088E-01 | 1.329E-01 | 1.250E-02 |
| C1 Chrysene c | 242 | 1.214E+00 | 1.434E+00 | 2.397E-02 | 4.517E-01 | 8.480E-01 | 1.561E+00 | 3.983E+00 | 7.506E-03 | 6.367E-03 | 5.285E-03 | 1.150E+00 | 1.936E-01 | 3.512E-02 |
| C1 Chrysene d | 242 | 1.155E+00 | 1.337E+00 | 3.862E-02 | 4.227E-01 | 8.480E-01 | 1.544E+00 | 3.904E+00 | 1.708E-02 | 1.480E-02 | 1.294E-02 | 5.481E+00 | 1.150E+00 | 3.631E-02 |
| Benzo[b]fluoranthene | 252 | 2.322E+00 | 6.474E+00 | 1.422E-01 | 1.831E+00 | 8.016E+00 | 3.140E+01 | 5.072E+01 | 4.546E-01 | 3.973E-01 | 3.472E-01 | 2.381E+00 | 4.503E-01 | 2.163E-02 |
| Benzo[k]fluoranthene | 252 | 1.289E+00 | 3.371E+00 | 8.899E-02 | 1.001E+00 | 4.044E+00 | 1.424E+01 | 2.331E+01 | 9.906E-01 | 7.141E-01 | 6.763E-01 | 4.583E+00 | 1.104E+00 | 6.684E-02 |
| benzofluoranthene c | 252 | 5.119E-01 | 1.251E+00 | 5.530E-02 | 3.754E-01 | 1.610E+00 | 5.071E+00 | 7.813E+00 | 6.145E-01 | 5.430E-01 | 4.475E-01 | 3.379E+00 | 7.938E-01 | 8.177E-02 |
| benzofluoranthene d | 252 | 9.026E-01 | 1.867E+00 | 4.057E-02 | 4.542E-01 | 1.366E+00 | 6.795E+00 | 7.038E+00 | 2.899E-01 | 2.611E-01 | 2.156E-01 | 1.132E+00 | 4.361E-01 | 2.491E-02 |
| Benzo[a]pyrene | 252 | 3.445E+00 | 7.821E+00 | 1.949E-01 | 1.719E+00 | 9.229E+00 | 2.814E+01 | 4.248E+01 | 1.045E+00 | 8.123E-01 | 6.184E-01 | 4.029E+00 | 1.020E+00 | 9.257E-02 |
| Benzopyrene b | 252 | 1.955E+00 | 4.202E+00 | 1.642E-01 | 1.012E+00 | 5.000E+00 | 1.944E+01 | 2.742E+01 | 9.238E-01 | 7.399E-01 | 6.953E-01 | 4.674E+00 | 1.033E+00 | 8.099E-02 |
| Indeno[1,2,3-cd]pyrene | 276 | 9.128E-01 | 2.587E+00 | 7.546E-02 | 6.406E-01 | 3.385E+00 | 1.511E+01 | 1.590E+01 | 7.046E-01 | 6.004E-01 | 4.838E-01 | 1.816E+00 | 5.887E-01 | 5.409E-02 |
| Dibenzo[a,h]anthracene | 278 | 1.602E-01 | 6.066E-01 | 1.704E-02 | 1.091E-01 | 4.620E-01 | 1.410E+00 | 4.381E+00 | 1.170E-01 | 1.033E-01 | 8.257E-02 | 6.349E-01 | 1.698E-01 | 1.196E-02 |
| Benzo[g,h,i]perylene | 276 | 6.862E-01 | 3.001E+00 | 1.094E-01 | 5.396E-01 | 3.417E+00 | 1.191E+01 | 1.655E+01 | 6.685E-01 | 5.363E-01 | 4.333E-01 | 1.285E+00 | 3.267E-01 | 4.770E-02 |
| dibenzochrysene | 276 | 3.973E-01 | 1.269E+00 | 2.907E-02 | 2.126E-01 | 6.764E-01 | 5.098E+00 | 5.489E+00 | 2.273E-01 | 2.145E-01 | 1.590E-01 | 4.338E-01 | 1.786E-01 | 2.016E-02 |
| 2-phenylnaphthalene | 204 | 6.869E+00 | 8.295E+00 | 5.947E-02 | 1.134E+00 | 3.522E+00 | 5.840E+00 | 1.071E+01 | 5.055E-01 | 4.502E-01 | 3.495E-01 | 2.272E+00 | 4.629E-01 | 6.023E-02 |
| phenalenothiophene a | 208 | 7.388E+00 | 1.387E+01 | 4.172E-02 | 4.728E-01 | 4.418E+00 | 4.304E+01 | 1.002E+01 | 5.101E-01 | 4.954E-01 | 3.684E-01 | 1.663E+00 | 2.258E-01 | 4.271E-01 |
| phenalenothiophene b | 208 | 1.062E+00 | 2.054E+00 | 1.374E-02 | 1.871E-01 | 1.199E+00 | 1.319E+00 | 1.711E+00 | 1.333E-01 | 1.234E-01 | 9.728E-02 | 2.085E+00 | 7.334E-02 | 9.166E-03 |
| anthracene, 9-ethenyl | 204 | 2.767E+00 | 2.237E+00 | 2.595E-02 | 3.980E-01 | 2.874E+00 | 4.920E-01 | 1.230E+01 | 1.819E-01 | 3.488E-01 | 2.682E-01 | 1.268E+00 | 9.928E-02 | 2.464E-02 |

| Compound | Mass (g/mol) | Samples | | | |
|---|---|---|---|---|---|
| | | P | D | PD | W19 |
| undecane (C11) | 156 | 1.854E-01 | 3.190E+00 | 7.758E+00 | 2.505E-07 |
| dodecane (C12) | 170 | 1.453E-01 | 3.201E+00 | 9.425E+00 | 2.505E-07 |
| tridecane (C13) | 184 | 3.301E-02 | 4.429E+00 | 1.087E+01 | 2.505E-07 |
| tetradecane (C14) | 198 | 3.397E-02 | 4.788E+00 | 1.141E+01 | 2.505E-07 |
| pentadecane (C15) | 212 | 3.087E-02 | 1.155E-01 | 1.397E-01 | 2.505E-07 |
| hexadecane (C16) | 226 | 2.230E-02 | 1.087E+01 | 1.504E+01 | 2.505E-07 |
| heptadecane (C17) | 240 | 1.804E-02 | 1.269E+01 | 1.578E+01 | 2.505E-07 |
| octadecane (C18) | 254 | 1.074E-02 | 1.571E+01 | 1.505E-01 | 2.505E-07 |
| tetracosane (C24) | 338 | 6.478E-04 | 4.959E+00 | 3.380E+00 | 2.505E-07 |
| pentacosane (C25) | 352 | 3.482E-04 | 2.496E+00 | 1.959E+00 | 2.505E-07 |
| hexacosane (C26) | 366 | 3.080E-04 | 1.351E+00 | 1.006E+00 | 2.505E-07 |
| heptacosane (C27) | 380 | 3.471E-04 | 5.905E-01 | 4.350E-01 | 2.505E-07 |
| octacosane (C28) | 394 | 2.931E-04 | 2.808E-01 | 2.825E-01 | 2.505E-07 |
| nonacosane (C29) | 408 | 2.504E-07 | 1.594E-01 | 1.304E-01 | 2.505E-07 |
| triacontane (C30) | 422 | 2.504E-07 | 9.729E-02 | 6.151E-02 | 2.505E-07 |
| hentriacontane (C31) | 436 | 2.504E-07 | 5.423E-02 | 4.169E-02 | 2.505E-07 |
| dotriacontane (C32) | 450 | 2.504E-07 | 2.370E-02 | 1.507E-02 | 2.505E-07 |
| tritriacontane (C33) | 464 | 2.504E-07 | 1.120E-02 | 8.305E-03 | 2.505E-07 |
| pristane | 170 | 1.432E-02 | 1.181E+01 | 1.048E+01 | 2.505E-07 |
| phytane | 184 | 9.200E-03 | 1.205E+01 | 1.118E+01 | 2.505E-07 |
| branched alkane A (C12) | 198 | 1.403E-01 | 8.180E+00 | 7.021E+00 | 2.505E-07 |
| branched alkane B (C13) | 268 | 8.485E-01 | 8.663E+00 | 8.976E+00 | 2.505E-07 |
| branched alkane C (C14) | 282 | 2.062E-02 | 9.468E+00 | 1.152E+01 | 2.505E-07 |
| phenol | 94 | 2.504E-07 | 2.502E-07 | 2.502E-07 | 2.030E-02 |
| C1 phenol a | 108 | 2.504E-07 | 2.502E-07 | 2.502E-07 | 2.975E-03 |
| C1 phenol b | 108 | 2.504E-07 | 2.502E-07 | 2.502E-07 | 1.201E-02 |
| C2 phenol a | 122 | 2.504E-07 | 2.502E-07 | 2.502E-07 | 1.330E-04 |
| C2 phenol b | 122 | 2.504E-07 | 2.502E-07 | 2.502E-07 | 3.137E-03 |
| C2 phenol c | 122 | 2.504E-07 | 2.502E-07 | 2.502E-07 | 4.233E-03 |
| C2 phenol d | 122 | 2.504E-07 | 2.502E-07 | 2.502E-07 | 5.569E-05 |
| toluene | 91 | 1.372E+00 | 5.423E-01 | 1.046E+00 | 2.505E-07 |
| C3-enyl benzene | 120 | 4.814E-01 | 1.174E+00 | 6.833E-01 | 1.396E-03 |
| C3 benzene a | 120 | 5.712E-01 | 1.503E+00 | 8.692E+00 | 1.470E-03 |
| C3 benzene b | 120 | 1.401E+00 | 2.869E+00 | 3.052E+00 | 9.169E-04 |
| C4 benzene a | 134 | 1.264E+00 | 7.369E-01 | 1.760E-01 | 2.105E-04 |
| C4 benzene b | 134 | 1.813E+00 | 3.567E-01 | 4.554E-01 | 2.505E-07 |
| C4 benzene c | 134 | 3.770E+00 | 7.303E-01 | 1.669E+00 | 2.505E-07 |
| C4 benzene d | 134 | 1.504E+00 | 9.032E-01 | 1.928E+00 | 7.493E-04 |
| C4 benzene e | 134 | 1.455E+00 | 7.769E-01 | 1.136E+00 | 2.505E-07 |
| C4 benzene f | 134 | 1.860E+00 | 5.946E-01 | 9.056E-01 | 2.006E-04 |
| C4 benzene g | 134 | 1.014E+00 | 4.788E-01 | 8.036E-01 | 2.505E-07 |

| Compound | Mass (g/mol) | Samples | | | |
|---|---|---|---|---|---|
| | | P | D | PD | W19 |
| indane | 118 | 1.143E+00 | 6.246E-01 | 1.167E+00 | 3.242E-04 |
| indene | 116 | 1.506E+00 | 2.502E-07 | 8.092E-01 | 4.428E-02 |
| benzofuran | 118 | 2.504E-07 | 2.502E-07 | 2.502E-07 | 1.470E-03 |
| benzothiophene | 134 | 2.504E-07 | 2.502E-07 | 2.502E-07 | 1.780E-02 |
| C1 benzothiophene a | 148 | 2.504E-07 | 2.502E-07 | 2.502E-07 | 7.650E-04 |
| C1 benzothiophene b | 148 | 2.504E-07 | 2.502E-07 | 2.502E-07 | 7.879E-04 |
| C1 benzothiophene c | 148 | 2.504E-07 | 2.502E-07 | 2.502E-07 | 6.640E-04 |
| C2 benzothiophene a | 162 | 2.504E-07 | 2.502E-07 | 2.502E-07 | 4.718E-05 |
| C2 benzothiophene b | 162 | 2.504E-07 | 2.502E-07 | 2.502E-07 | 1.619E-04 |
| C2 benzothiophene c | 162 | 2.504E-07 | 2.502E-07 | 2.502E-07 | 1.026E-04 |
| C2 benzothiophene d | 162 | 2.504E-07 | 2.502E-07 | 2.502E-07 | 1.230E-04 |
| C2 benzothiophene e | 162 | 2.504E-07 | 2.502E-07 | 2.502E-07 | 8.633E-05 |
| C2 benzothiophene f | 162 | 2.504E-07 | 2.502E-07 | 2.502E-07 | 8.705E-05 |
| C2 benzothiophene g | 162 | 2.504E-07 | 2.502E-07 | 2.502E-07 | 2.505E-07 |
| C2 benzothiophene h | 162 | 2.504E-07 | 2.502E-07 | 2.502E-07 | 2.505E-07 |
| C2 benzothiophene i | 162 | 2.504E-07 | 2.502E-07 | 2.502E-07 | 2.505E-07 |
| C3 benzothiophene a | 176 | 2.504E-07 | 2.502E-07 | 2.502E-07 | 2.505E-07 |
| C3 benzothiophene b | 176 | 2.504E-07 | 2.502E-07 | 2.502E-07 | 2.505E-07 |
| C3 benzothiophene c | 176 | 2.504E-07 | 2.502E-07 | 2.502E-07 | 2.505E-07 |
| C3 benzothiophene d | 176 | 2.504E-07 | 2.502E-07 | 2.502E-07 | 2.505E-07 |
| C3 benzothiophene e | 176 | 2.504E-07 | 2.502E-07 | 2.502E-07 | 2.505E-07 |
| diphenylmethane | 167 | 4.726E-03 | 6.957E-02 | 8.558E-02 | 6.542E-05 |
| 1,2-dihydronaphthalene | 130 | 1.358E-01 | 2.502E-07 | 6.797E-02 | 2.505E-07 |
| 1,4-dihydronaphthalene | 130 | 8.481E-02 | 2.502E-07 | 3.172E-02 | 2.505E-07 |
| tetrahydronaphthalene | 132 | 4.053E-02 | 7.201E-01 | 8.417E-01 | 2.505E-07 |
| naphthalene | 128 | 2.775E+00 | 7.422E-01 | 4.185E+00 | 2.365E-01 |
| 2-methyl naphthalene | 142 | 1.232E+00 | 1.426E+00 | 3.063E+00 | 6.025E-02 |
| 1-methyl naphthalene | 142 | 6.456E-01 | 1.129E+00 | 2.048E+00 | 3.413E-02 |
| 2-EtN | 156 | 6.093E-02 | 4.190E-01 | 3.905E-01 | 4.981E-04 |
| 2,6-DMN & 2,7-DMN | 156 | 1.405E-01 | 9.713E-01 | 1.306E+00 | 7.807E-03 |
| 1-EtN | 156 | 3.123E-02 | 7.898E-02 | 2.477E-01 | 2.281E-03 |
| 1,6-DMN | 156 | 7.882E-02 | 9.121E-01 | 1.247E+00 | 3.076E-03 |
| 1,7-DMN & 1,3-DMN | 156 | 1.633E-01 | 1.099E+00 | 1.533E+00 | 1.027E-02 |
| 2,3-DMN | 156 | 3.450E-02 | 1.863E-01 | 2.742E-01 | 2.111E-03 |
| 1,4-DMN | 156 | 3.005E-02 | 1.930E-01 | 2.067E-01 | 1.561E-03 |
| 1,5-DMN | 156 | 1.677E-02 | 2.707E-01 | 3.357E-01 | 1.604E-03 |
| 1,2-DMN | 156 | 4.878E-01 | 2.397E-01 | 3.655E-01 | 1.528E-03 |
| 1,8-DMN | 156 | 2.504E-07 | 2.502E-07 | 2.502E-07 | 2.505E-07 |
| C3 alkyl naphthalene a | 170 | 1.807E-02 | 4.327E-01 | 4.699E-01 | 2.505E-07 |
| C3 alkyl naphthalene b | 170 | 3.846E-02 | 7.643E-01 | 5.858E-01 | 5.092E-04 |
| C3 alkyl naphthalene c | 170 | 1.301E-02 | 9.351E-01 | 7.318E-01 | 2.978E-03 |

Table C.13: Normalised GCxGC peak areas for all coal tar samples - raw data 10/10

| Sample | Mass (g/mol) | P | D | PD | W19 |
|---|---|---|---|---|---|
| C3 alkyl naphthalene d | 170 | 4.359E-02 | 1.370E+00 | 1.301E+00 | 2.407E-03 |
| C3 alkyl naphthalene e | 170 | 6.415E-02 | 8.567E-01 | 1.385E+00 | 1.431E-03 |
| C3 alkyl naphthalene f | 170 | 3.720E-02 | 6.022E-01 | 1.075E+00 | 4.694E-04 |
| C3 alkyl naphthalene g | 170 | 2.262E-02 | 5.532E-02 | 5.816E-01 | 8.317E-05 |
| C3 alkyl naphthalene h | 170 | 9.997E-03 | 2.090E-01 | 3.450E-01 | 2.333E-04 |
| C3 alkyl naphthalene i | 170 | 8.651E-03 | 3.120E-02 | 2.466E-01 | 5.551E-05 |
| biphenyl | 154 | 1.933E-02 | 1.086E+00 | 1.255E+00 | 1.906E-02 |
| C1 biphenyl a | 168 | 2.865E-02 | 9.441E-01 | 1.229E+00 | 4.446E-03 |
| C1 biphenyl b | 168 | 1.613E-02 | 2.520E-01 | 3.857E-01 | 2.374E-03 |
| dibenzofuran | 168 | 1.365E-03 | 1.014E-01 | 1.145E-01 | 1.255E-01 |
| C1 dibenzofuran a | 182 | 2.504E-07 | 2.502E-07 | 2.502E-07 | 1.078E-02 |
| C1 dibenzofuran b | 182 | 2.504E-07 | 2.502E-07 | 2.502E-07 | 1.408E-02 |
| C1 dibenzofuran c | 182 | 2.504E-07 | 2.502E-07 | 2.502E-07 | 3.547E-03 |
| acenaphthylene | 152 | 2.550E-03 | 1.583E-03 | 6.085E-03 | 2.772E-01 |
| acenaphthene | 154 | 8.029E-03 | 4.771E-02 | 7.115E-02 | 1.056E-02 |
| dibenzothiophene | 184 | 2.504E-07 | 2.502E-07 | 2.502E-07 | 2.263E-02 |
| naphtho-thiophene | 184 | 2.504E-07 | 2.502E-07 | 2.502E-07 | 1.269E-02 |
| C1 dibenzothiophene a | 198 | 2.504E-07 | 2.502E-07 | 2.502E-07 | 1.412E-03 |
| C1 dibenzothiophene b | 198 | 2.504E-07 | 2.502E-07 | 2.502E-07 | 1.047E-03 |
| C1 dibenzothiophene c | 198 | 2.504E-07 | 2.502E-07 | 2.502E-07 | 1.138E-03 |
| C2 dibenzothiophene a | 212 | 2.504E-07 | 2.502E-07 | 2.502E-07 | 1.062E-04 |
| C2 dibenzothiophene b | 212 | 2.504E-07 | 2.502E-07 | 2.502E-07 | 1.292E-04 |
| C2 dibenzothiophene c | 212 | 2.504E-07 | 2.502E-07 | 2.502E-07 | 1.991E-04 |
| C2 dibenzothiophene d | 212 | 2.504E-07 | 2.502E-07 | 2.502E-07 | 2.346E-04 |
| C2 dibenzothiophene e | 212 | 2.504E-07 | 2.502E-07 | 2.502E-07 | 1.147E-04 |
| C2 dibenzothiophene f | 212 | 2.504E-07 | 2.502E-07 | 2.502E-07 | 1.709E-04 |
| C3 dibenzothiophene a | 226 | 2.504E-07 | 2.502E-07 | 2.502E-07 | 5.825E-05 |
| C3 dibenzothiophene b | 226 | 2.504E-07 | 2.502E-07 | 2.502E-07 | 4.193E-05 |
| C3 dibenzothiophene c | 226 | 2.504E-07 | 2.502E-07 | 2.502E-07 | 2.253E-05 |
| C3 dibenzothiophene d | 226 | 2.504E-07 | 2.502E-07 | 2.502E-07 | 1.269E-05 |
| C3 dibenzothiophene e | 226 | 2.504E-07 | 2.502E-07 | 2.502E-07 | 1.798E-05 |
| C3 dibenzothiophene f | 226 | 2.504E-07 | 2.502E-07 | 2.502E-07 | 3.465E-05 |
| C3 dibenzothiophene g | 226 | 2.504E-07 | 2.502E-07 | 2.502E-07 | 4.235E-05 |
| C3 dibenzothiophene h | 226 | 2.504E-07 | 2.502E-07 | 2.502E-07 | 2.257E-05 |
| fluorene | 166 | 2.844E-02 | 2.700E-01 | 3.136E-01 | 1.254E-01 |
| C1 fluorene a | 180 | 4.334E-02 | 2.318E-01 | 2.790E-01 | 6.893E-03 |
| C1 fluorene b | 180 | 4.495E-02 | 3.675E-01 | 4.520E-01 | 5.161E-03 |

| Compound | Mass (g/mol) | P | D | PD | W19 |
|---|---|---|---|---|---|
| phenanthrene | 178 | 4.224E-02 | 4.814E-01 | 5.351E-01 | 2.906E-01 |
| anthracene | 178 | 1.063E-02 | 3.364E-02 | 3.494E-02 | 1.107E-01 |
| C1 phenanthrene a | 192 | 1.602E-02 | 3.100E-01 | 3.460E-01 | 1.174E-02 |
| C1 phenanthrene b | 192 | 1.328E-02 | 4.444E-01 | 4.241E-01 | 6.591E-03 |
| C1 phenanthrene c | 192 | 1.526E-02 | 2.502E-07 | 2.502E-07 | 2.070E-02 |
| C1 anthracene a | 192 | 5.017E-03 | 3.154E-01 | 3.154E-01 | 3.261E-03 |
| C1 anthracene b | 192 | 9.224E-03 | 3.504E-01 | 3.367E-01 | 4.459E-03 |
| C1 anthracene c | 192 | 2.504E-07 | 2.502E-07 | 2.502E-07 | 8.248E-03 |
| Carbazole | 167 | 2.504E-07 | 2.502E-07 | 2.502E-07 | 8.866E-02 |
| Fluoranthene | 202 | 5.772E-03 | 1.573E-02 | 1.368E-02 | 1.922E-01 |
| Pyrene | 202 | 1.058E-02 | 1.644E-01 | 1.714E-01 | 1.346E-01 |
| C1 fluoranthene/pyrene a | 216 | 2.405E-03 | 1.716E-02 | 2.036E-02 | 5.353E-03 |
| C1 fluoranthene/pyrene b | 216 | 2.397E-03 | 3.834E-03 | 2.393E-02 | 2.818E-02 |
| C1 fluoranthene/pyrene c | 216 | 1.147E-03 | 4.851E-03 | 3.758E-02 | 1.852E-02 |
| C1 fluoranthene/pyrene d | 216 | 3.602E-03 | 2.608E-02 | 2.034E-02 | 1.324E-02 |
| C1 fluoranthene/pyrene e | 216 | 1.595E-03 | 9.219E-02 | 8.100E-02 | 6.837E-03 |
| C1 fluoranthene/pyrene f | 216 | 1.276E-03 | 1.468E-01 | 1.396E-01 | 6.082E-03 |
| Benz[a]anthracene | 228 | 1.025E-03 | 1.109E-03 | 2.278E-03 | 6.036E-02 |
| Chrysene | 228 | 4.370E-03 | 1.410E-02 | 1.627E-02 | 8.951E-02 |
| C1 Chrysene a | 242 | 2.504E-07 | 5.163E-04 | 6.096E-03 | 3.675E-04 |
| C1 Chrysene b | 242 | 5.820E-04 | 2.725E-03 | 3.549E-03 | 3.462E-04 |
| C1 Chrysene c | 242 | 2.504E-07 | 7.861E-04 | 9.629E-04 | 4.372E-03 |
| C1 Chrysene d | 242 | 2.590E-04 | 4.638E-03 | 5.409E-03 | 4.721E-03 |
| Benzo[b]fluoranthene | 252 | 6.843E-04 | 1.513E-03 | 2.584E-03 | 5.364E-02 |
| Benzo[k]fluoranthene | 252 | 5.910E-04 | 1.206E-03 | 2.328E-03 | 5.782E-02 |
| benzofluoranthene c | 252 | 4.751E-04 | 4.767E-04 | 2.334E-03 | 3.637E-02 |
| benzofluoranthene d | 252 | 4.485E-04 | 5.773E-04 | 2.595E-03 | 2.758E-02 |
| Benzo[a]pyrene | 252 | 1.164E-03 | 1.242E-03 | 2.538E-03 | 5.353E-02 |
| Benzopyrene b | 252 | 8.088E-04 | 8.482E-04 | 4.111E-03 | 5.481E-02 |
| Indeno[1,2,3-cd]pyrene | 276 | 7.800E-04 | 5.972E-04 | 3.848E-03 | 3.917E-02 |
| Dibenzo[a,h]anthracene | 278 | 1.041E-04 | 1.217E-04 | 5.256E-04 | 3.963E-03 |
| Benzo[g,h,i]perylene | 276 | 1.144E-03 | 1.070E-03 | 5.944E-03 | 4.355E-02 |
| dibenzochrysene | 226 | 2.863E-04 | 1.687E-04 | 2.071E-04 | 1.317E-02 |
| 2-phenylnaphthalene | 204 | 2.504E-07 | 2.502E-07 | 2.502E-07 | 8.084E-03 |
| phenalenothiophene a | 208 | 2.504E-07 | 2.502E-07 | 2.502E-07 | 6.084E-03 |
| phenalenothiophene b | 208 | 2.504E-07 | 2.502E-07 | 2.502E-07 | 1.310E-03 |
| anthracene, 9-ethenyl | 204 | 2.504E-07 | 2.502E-07 | 2.502E-07 | 4.424E-03 |

## C.4 Weathering Ratios

Table C.14: Weathering ratios calcuated for coal tars and petrogenics; as used in Figure 7.23.

| Sample | N/(N2+N3) | BT/(BT2+BT3) | DBT/(DBT2+DBT3) |
|--------|-----------|--------------|-----------------|
| 1 | 5.27 | 2.76 | 5.97 |
| 2 | 4.36 | 2.31 | 4.00 |
| 3 | 2.51 | 2.30 | 5.09 |
| 4 | 3.02 | 3.06 | 5.49 |
| 5 | 2.17 | 2.44 | 4.81 |
| 6 | 2.22 | 3.22 | 5.40 |
| 7 | 19.85 | 36.81 | 30.75 |
| 8 | 8.29 | 8.31 | 3.96 |
| 9 | 9.16 | 14.68 | 22.84 |
| 10 | 35.22 | 72.11 | 43.53 |
| 11 | 0.94 | 2.74 | 9.30 |
| 12 | 0.14 | 0.21 | 1.87 |
| 13 | 0.37 | 1.15 | 2.99 |
| 14 | 0.56 | 1.22 | 3.99 |
| 15 | 1.92 | 2.03 | 5.16 |
| 16 | 2.44 | 8.33 | 12.91 |
| 17 | 1.70 | 5.25 | 9.68 |
| 18 | 16.48 | 103.70 | 53.25 |
| 19 | 6.19 | 42.75 | 31.22 |
| 20 T1 | 0.41 | 2.61 | 9.36 |
| 20 T3 | 0.71 | 2.33 | 7.20 |
| 20 T5 | 0.76 | 2.71 | 7.40 |
| 21 | 0.77 | 1.41 | 6.16 |
| 22 | 1.93 | 9.20 | 4.84 |
| 23 | 1.67 | 2.28 | 2.19 |
| P | 3.22 | 0.07 | 0.07 |
| D | 0.08 | 0.07 | 0.07 |
| PD | 0.33 | 0.07 | 0.07 |
| W19 | 6.08 | 29.18 | 18.74 |

# Appendix D

# Peer-reviewed Publications

1. L. A. McGregor, C. Gauchotte-Lindsay, N. Nic Daéid, R. Thomas, P. Daly, R. M. Kalin, Ultra Resolution Chemical Fingerprinting of Dense Non-Aqueous Phase Liquids from Manufactured Gas Plants by Reversed Phase Comprehensive Two-Dimensional Gas Chromatography, *J. Chromatogr. A*, 2011, 1218, 4755-4763.

# Ultra resolution chemical fingerprinting of dense non-aqueous phase liquids from manufactured gas plants by reversed phase comprehensive two-dimensional gas chromatography

Laura A. McGregor [a,*], Caroline Gauchotte-Lindsay [a], Niamh Nic Daéid [b], Russell Thomas [c], Paddy Daly [d], Robert M. Kalin [a]

[a] Department of Civil Engineering, University of Strathclyde, John Anderson Building, 107 Rottenrow, Glasgow G4 0NG, UK
[b] Centre for Forensic Science, Department of Pure and Applied Chemistry, University of Strathclyde, Royal College Building, 204 George Street, Glasgow, UK
[c] Parsons Brinckerhoff, Queen Victoria House, Redland Hill, Bristol, UK
[d] National Grid, National Grid House, Warwick Technology Park, Gallows Hill, Warwick, UK

## ARTICLE INFO

## ABSTRACT

Ultra resolution chemical fingerprinting of dense non-aqueous phase liquids (DNAPLs) from former manufactured gas plants (FMGPs) was investigated using comprehensive two-dimensional gas chromatography coupled with time of flight mass spectrometry (GC × GC TOFMS). Reversed phase GC × GC (i.e. a polar primary column coupled to a non-polar secondary column) was found to significantly improve the separation of polycyclic aromatic hydrocarbons (PAHs) and their alkylated homologues. Sample extraction and cleanup was performed simultaneously using accelerated solvent extraction (ASE), with recovery rates between 76% and 97%, allowing fast, efficient extraction with minimal solvent consumption. Principal component analysis (PCA) of the GC × GC data was performed in an attempt to differentiate between twelve DNAPLs based on their chemical composition. Correlations were discovered between DNAPL composition and historic manufacturing processes used at different FMGP sites. Traditional chemical fingerprinting methods generally follow a tiered approach with sample analysis on several different instruments. We propose ultra resolution chemical fingerprinting as a fast, accurate and precise method of obtaining more chemical information than traditional tiered approaches while using only a single analytical technique.

© 2011 Elsevier B.V. All rights reserved.

## 1. Introduction

A dense non-aqueous phase liquid (DNAPL) is a liquid which is both heavier than water and immiscible in water [1]. In this case, DNAPL refers to coal tar; a common subsurface contaminant found at former manufactured gas plants (FMGPs). Coal tar DNAPLs are composed of thousands of organic and inorganic compounds, many of which may be found in trace quantities [2].

The complex chemical composition of DNAPLs has been shown to vary dramatically within a single FMGP site, as well as between different sites [3]. Accurate chemical fingerprinting is required at FMGP sites to ensure multiple sources of contamination are not present [4]. For example, more recent spills could be distinguished from historical gasworks contamination. Furthermore, for FMGPs split into multiple land holdings, accurate chemical fingerprinting can help to identify liability across the entire site. Given the large

number of former gasworks sites in the U.K. and the introduction of recent "polluters pay" legislation, [5] it is reasonable to assume there may be many liability cases in the future, thus spurring the growth of the environmental forensics industry in the U.K.

Environmental forensic chemical fingerprinting of complex samples, such as coal tar and crude oil, is generally performed by gas chromatography (GC) in combination with either flame ionisation detection (GC–FID) or mass spectrometry (GC–MS) within a tiered analytical approach [4,6–8]. However, conventional GC techniques do not have the capacity to resolve the complex composition of coal tar DNAPLs [9]. Time-consuming and labour-intensive chemical fractionation processes are generally required to divide complex mixtures into several extracts prior to analysis [10].

There have been few reports on DNAPL composition in recent literature [3,11,12] and to the authors' knowledge there is no standardised approach for analysis of free phase coal tars, certainly not without extensive sample fractionation. Brown et al. [3] evaluated the composition of DNAPLs from ten different FMGP sites in the U.S.A indicating major differences in PAH composition between sites. However, this study utilised GC–MS analysis after lengthy

fractionation processes, so the chemical information obtained on the DNAPLs was limited by resolution power of the technique. Generally, the literature focuses on challenges involved in characterisation and remediation of DNAPL contaminated land [2,13–15]. For example, Birak and Miller [2] state that full characterisation of DNAPLs at FMGP sites is still limited by analytical techniques. Utilisation of advanced chromatographic techniques for chemical fingerprinting of DNAPLs has thus been long-awaited to aid characterisation and allow the most effective remediation routes to be chosen.

Comprehensive two-dimensional gas chromatography (GC × GC) is a high-resolution separation technique, developed with the intention of overcoming limitations associated with conventional GC techniques [16]. The coupling of two columns with different selectivity allows for a two-dimensional separation of mixtures, across a retention plane rather than along a retention line [17–21]. An order of magnitude more compounds can be separated by GC × GC than when using conventional GC instrumentation [22].

Generally, a long, wide bore (0.25–0.32 mm i.d.), non-polar capillary column is used in the first separation, whereas a short, narrow bore (0.1–0.2 mm i.d.), polar column is installed for the second separation; this is deemed normal phase. However, reversing the column polarity has been shown to provide better group-type separation in certain cases [23]. The use of a polar, primary column and non-polar, secondary column is known as reversed phase (or reversed polarity) GC × GC [17].

GC × GC has been shown to be especially useful for environmental forensic analyses of complex samples [24,25]; the main advantage being the minimisation or elimination of fractionation processes prior to analysis [16,26]. A complex sample can be injected as a single extract to provide fast screening of the entire sample, allowing many classes of organic contaminants to be monitored at once. However, the technique has yet to be applied to the analysis of free phase coal tars.

This work aims to use GC × GC TOFMS to resolve the issues associated with the analysis and source apportionment of coal tar DNAPLs. Chemical fingerprinting of environmental samples by conventional GC techniques is described as a high resolution method. In this study, we demonstrate an enhanced method of chemical fingerprinting, deemed 'ultra resolution', by combining reversed phase GC × GC with statistical comparison using principal components analysis (PCA). This process gathers more chemical information per sample than traditional tiered approaches and has the additional benefits of using an efficient one-step extraction followed by analysis on a single analytical instrument.

## 2. Experimental

### 2.1. Samples and standards

DNAPL samples (labelled 1–12) were provided from seven different FMGP sites across the United Kingdom. The gas manufacturing processes used at each site are summarised in Table 1. DNAPL samples 1–6 were obtained from various locations within the same site (site A), while all other samples were acquired from different sites. Samples 1–10 were all obtained from sites that used coal retort stands for gas production, whereas sample 11 was obtained from a wood preservative site, where coal tar was distilled to produce creosote oil for coating wood [27]. Sample 12 was obtained from a carburetted water gas (CWG) plant where a mixture of hydrogen and carbon monoxide was produced by passing steam through heated coke rather than by the carbonisation of coal performed at retort gasworks [28]. The samples were stored at 4 °C prior to analysis.

**Table 1**
Description of FMGP sites.

| Sample no. | Manufacturing process(es) | Sampling location |
|---|---|---|
| 1 | Vertical coal retort; potential traces of horizontal retort tar and gas oil (from micro-simplex gas reforming plant on site) | Borehole |
| 2 | " | Borehole near gas holder |
| 3 | " | Within tar tank |
| 4 | " | Within tar tank |
| 5 | " | Within tar tank |
| 6 | " | Borehole near tar tank |
| 7 | Horizontal coal retort | Base of gas holder |
| 8 | Horizontal coal retort | Within tar tank |
| 9 | Vertical coal retort; potential traces of carburetted water gas tar and horizontal retort tar | Unknown |
| 10 | Horizontal coal retort | Unknown |
| 11 | Wood preservation site; tar probably from a distilled fraction of creosote oil | Sump |
| 12 | Complex mixture of horizontal and vertical retorts, water gas and gas oil (from a gas reforming plant on site) | Borehole |

All solvents used (*n*-hexane, dichloromethane) were of analytical grade, purchased from Fisher Scientific (Loughborough, U.K) and used without further purification. All deuterated PAHs were obtained from Isotec™, Sigma–Aldrich (Gillingham, U.K). All PAHs and alkylated naphthalenes were purchased from Sigma–Aldrich.

Anhydrous sodium sulphate, silica gel 60 (both from Sigma–Aldrich) and diatomaceous earth (Dionex, Camberley, UK) were activated for 4 h at 450 °C prior to use. Silica gel 60 was then deactivated by 10% water (w/w).

Alkylated naphthalenes were identified in the DNAPL extracts using individually prepared 200 μg/mL (in dichloromethane) standards of 1- and 2-methyl naphthalene and the 12 C2 alkyl naphthalene isomers.

Target analytes in the DNAPL extracts were quantified using calibration mixtures containing 16 PAHs, priority pollutants as listed by the U.S. EPA [29]. The 16 PAHs were purchased as a 2000 μg/mL stock solution in benzene:dichloromethane (1:1) from Sigma–Aldrich (Gillingham, U.K). A 2000 μg/mL stock surrogate solution containing deuterated PAHs (D8-naphthalene, D10-fluorene, D10-fluoranthene and D12-chrysene) was prepared to monitor extraction efficiency. Seven calibration standards containing the PAHs and surrogates were prepared within the concentration range of 2.5–500 μg/mL, each spiked with 75 μL of a 2000 μg/mL stock solution of D10-phenanthrene as an internal standard. Quantification was performed using the response of specific target ions present in GC × GC chromatograms (target ions are listed in Table S1 of supplementary data).

### 2.2. Sample preparation

Extraction was performed using an ASE 350 Accelerated Solvent Extraction system (Dionex, Camberley, UK) equipped with 10 mL stainless steel extraction cells. The high separation capability of GC × GC TOFMS eliminates the requirement for sample fractionation, thus a single extraction using hexane (including in-cell cleanup by silica gel) was performed.

A dry, homogeneous mix of DNAPL was prepared by grinding the DNAPL (approximately 0.5 g) with sodium sulphate (NaSO₄) and diatomaceous earth (D.E.) in a 1:1:1 ratio. This removes any water present in the DNAPL sample and results in a fine powder (rather than a tar) which can be transferred quantitatively to the

extraction cells. To ensure accurate quantification, the DNAPL was spiked with 600 µL of the surrogate solution prior to grinding with D.E. and NaSO$_4$. Any loss of target analytes could then be monitored from the start of sample preparation and storage of the sample in this form also allows any loss of target analytes over time to be monitored.

Extraction cells were lined with 2 filter papers (to ensure unwanted particulate matter did not collect in the extract) and packed with 3 g silica gel 60 (10% deactivated w/w). Approximately 0.5 g of the ground DNAPL/surrogate mixture was added to the extraction cell and the remaining cell volume was packed with D.E. Hexane was used as the extracting solvent for all extractions. ASE was performed at 150 °C and 10 MPa, using one dynamic (7 min) and two static (5 min each) extractions. A flush volume of 150% and purge time of 60 s were used. The extracts were concentrated to 1 mL using a Büchi Syncore® Analyst (Oldham, U.K). The extracts were then made up to exactly 10 mL using hexane. A 1 mL aliquot was then transferred to an autosampler vial and spiked with 75 µL of internal standard prior to analysis.

## 2.3. GC–MS analyses

A Thermo Scientific (Hertfordshire, U.K.) Trace Ultra GC fitted with a DSQII mass spectrometer and Triplus autosampler was used for all GC–MS analyses. The column was a J&W Scientific DB-5 fused silica capillary column (30 m × 0.25 mm i.d. × 0.25 µm film thickness). All injections were of one microlitre and were carried out using a split ratio of 1:50 and injection port temperature of 230 °C. Helium was used as the carrier gas, with a flow rate of 1.0 mL/min. All standards and extracts were analysed with the oven temperature programmed at 10 °C/min from 55 °C (maintained for 2 min) to 110 °C, 3 °C/min to 210 °C, then at 8 °C/min to 320 °C (maintained for 15 min).

## 2.4. GC × GC TOFMS analyses

All GC × GC TOFMS analyses were performed using a Leco (St. Joseph, Michigan) time of flight mass spectrometer, model Pegasus 4D, connected to an Agilent 7890A gas chromatograph equipped with a Leco thermal modulator. The TOF ion source was fixed at 200 °C and masses between 45 and 500u were scanned at a 200 spectra/second rate. The detector voltage was set at 1700 V and the applied electron ionisation voltage was set at 70 eV.

All standards and extracts were analysed with the primary oven temperature programmed at 10 °C/min from 55 °C (maintained for 2 min) to 110 °C, 3 °C/min to 210 °C, then at 8 °C/min to 310 °C (maintained for 15 min). The secondary oven and modulator temperatures were programmed at a 20 °C offset relative to the primary oven. The modulation period was 6 s with a 1.3 s hot pulse time. The injection port temperature was set to 250 °C using a split ratio of 1:50. One microlitre of sample was injected for each run using an MPS2 twister autosampler (Gerstel). Helium was used as the carrier gas, with a flow rate of 1.0 mL/min.

The normal phase column set comprised of a non-polar Rxi 5-Sil MS (25 m × 0.25 mm i.d. × 0.25 µm film thickness) primary column coupled to a mid-polarity Rxi 17 (1.2 m × 0.1 mm i.d. × 0.1 µm film thickness) secondary column, both supplied by Thames Restek (Buckinghamshire, U.K.). The reversed polarity column set comprised a mid-polarity TR-50 MS supplied by Thermo Scientific (30 m × 0.25 mm i.d. × 0.25 µm film thickness) as the primary column and a non-polar Rtx-5 supplied by Thames Restek (1.2 m × 0.18 mm i.d. m × 0.2 µm film thickness) as the secondary column, connected via a Thames Restek Press-tight® connector.

## 2.5. Principal component analysis

Variations in the DNAPL composition were evaluated by principal component analysis (PCA) using Minitab® 15 (Minitab Ltd., Coventry) software. Principal component analysis is a method used to extract the variations within a large data set by reducing raw sample data into smaller, uncorrelated variables known as principle components [30,31]. Score plots of the principal components which describe the most variation within the data allow relationships between the samples to be evaluated.

Peak areas of the tentatively identified compounds were imported into the statistical software after normalisation, against the peak area of the internal standard, and correction using the exact weight of DNAPL extracted for each sample.

## 3. Results and discussion

### 3.1. Optimisation of extraction procedure

The initial part of this study was dedicated to optimisation of extraction procedure, with the aim of extracting all chemical classes present in the DNAPL using a single Accelerated Solvent Extraction (ASE) method. Hexane was found to be a suitable extraction solvent, thus eliminating the need for harmful, chlorinated solvents. The ASE procedure utilised simultaneous extraction and clean-up, by the addition of silica gel to each extraction cell, thus further reducing the total analysis time and solvent consumption.

Fractionation of contaminated soil samples by ASE has previously been achieved by three separate extractions per cell using solvents of increasing polarity [32]. However, this was not possible for the DNAPL samples investigated in this study, as they were fully extracted by the initial, non-polar solvent despite attempts using low temperatures (40 °C) for the first extraction. GC–MS analysis of such complex samples would generally only be performed after chemical fractionation; however, given the ease of dissolution of the DNAPLs it is unlikely that effective fractionation could be achieved via ASE without the use of additional column chromatography. The high resolution capacity of GC × GC negates the requirement for sample fractionation thus the combination of sample extraction and cleanup by ASE provides fast screening of the entire coal tar composition.

Repeatability of the method was measured by extraction of six replicate cells, and subsequent GC–MS analysis, of DNAPL sample 7. Due to the difficulties involved in replicating a blank coal tar matrix, the surrogate recovery values were used as a measure of repeatability. Four deuterated PAHs (D8-naphthalene, D10-fluorene, D10-fluoranthene and D12-chrysene) were chosen as they span a range of molecular masses, from 136 g/mol to 240 g/mol. Recoveries between 76 and 97% were obtained based on the deuterated surrogate spikes. These values fall within the accepted range of 70–130% as stated by the U.S. EPA SW-846 Method 8000B [33]. Re-extraction of sample cells confirmed that the method provided exhaustive extraction of the DNAPL, with only the internal standard peak evident in the chromatograms of the second extracts. The relative standard deviation (RSD) of surrogate recovery was found to be below 10% for all deuterated surrogates, indicating satisfactory extraction repeatability.

### 3.2. Reversed polarity GC × GC

The column sets and GC × GC parameters were adjusted to achieve best possible separation of DNAPL components. Normal phase column sets are generally used in GC × GC analysis of environmental samples. Due to the restrictions in maximum operating temperature of most polar columns, a compromise generally exists
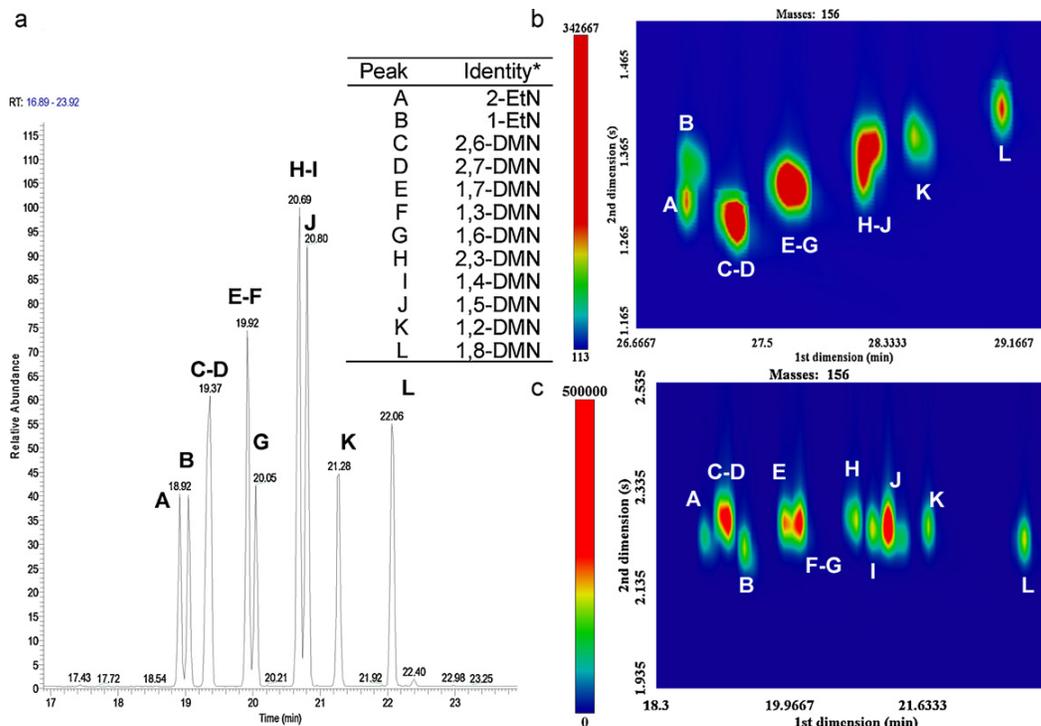
**Fig. 1.** Comparison of the separation capabilities of (a) GC–MS, (b) normal phase GC × GC and (c) reversed phase GC × GC using a standard mixture of C2 alkyl naphthalene isomers. The table provides the peak identities of the naphthalenes in each figure (*EtN = ethyl naphthalene, DMN = dimethyl naphthalene).

between column polarity and temperature programme for the secondary oven. The column phase was reversed to allow elution of the high molecular weight compounds present in DNAPLs while remaining within the limits of the column temperature range.

A standard mixture of C2 alkyl naphthalenes was used to confirm elution order and compare the separating power of three GC methods; GC–MS, normal phase GC × GC and reversed phase GC × GC. The C2 alkyl naphthalenes were chosen for this study as alkyl PAHs are often used in diagnostic ratios for source determination [7]. Due to insufficient separation with conventional GC techniques, the alkyl PAHs are generally combined by alkylation level, to provide diagnostic ratios based on quantification values for the group as a whole. For example, a typical diagnostic ratio using alkyl naphthalenes would be C0N/(C2N + C3N), where C0N is naphthalene and C2N and C3N are the C2 and C3 alkyl naphthalenes respectively [7]. We propose that the higher resolution of reversed phase GC × GC could allow enhanced diagnostic ratios to be calculated at no extra cost compared to normal phase GC × GC.

The chromatograms of the separation of a mixture of alkyl naphthalene isomers using GC–MS, normal phase GC × GC and reversed phase GC × GC are presented in Fig. 1. The GC × GC chromatograms are represented as contour plots; the x-axis represents the retention time in the primary column, the y-axis represents the retention time in the second column and the colour gradient represents the intensity of the peak. Normal phase GC × GC and GC–MS achieved separation of 7 and 9 peaks respectively. Reversed phase GC × GC allows separation of the 12 C2 alkyl naphthalenes into 10 peaks, with only 2 pairs of the alkyl naphthalenes still co-eluting (2,6- and 2,7-dimethyl naphthalene and 1,3- and 1,6-dimethyl naphthalene). Interestingly, normal phase GC × GC, which is generally used for the separation of complex samples, showed lower resolution for the alkyl naphthalenes than GC–MS.

The enhanced separation of reversed phase over normal phase GC × GC is further illustrated by chromatograms of the C3 and C4

alkyl naphthalenes (C3N and C4N respectively) in Fig. 2. Normal phase separates 9 peaks out of 34 possible C3N isomers and 14 of the 112 possible C4N isomers, while reversed phase separates 14 C3N and 20 C4N peaks within the same DNAPL sample. Full total ion chromatograms of DNAPL sample 1 by normal phase and reversed phase GC × GC TOFMS can be found in the supplementary data (Figs. S2 and S3 respectively).

The increased separation capacity of reversed phase is not only limited to alkyl PAHs. The DNAPL samples investigated in this study were found to contain a wide variety of chemical classes, including a range of alkylated heterocyclic PAH compounds. For example, alkyl benzothiophenes were abundant in all DNAPL samples. A comparison of the separating power of the two GC × GC modes for the benzothiophenes is shown in Fig. 3. The numbering indicates the peaks identified as alkyl benzothiophenes by their mass spectra, as some low intensity peaks can often be masked in the contour plot. Figs. 2 and 3 also illustrate the ordered structure of GC × GC contour plots; chemical families elute together in a band, allowing straightforward identification.

For example, the C1 alkyl naphthalenes elute together on a line with the higher alkylated homologues in subsequent bands. The structured layout of the contour plot allows peaks to be assigned quickly without the use of individual standards [34]. This form of tentative identification was used to assign the major chemical classes in the chromatogram of DNAPL 12 (Fig. 4) where the greatest variety of components was observed.

The 16 U.S. EPA priority PAHs are identified in Fig. 4, as well as their alkylated homologues. The elution order using reversed phase GC × GC is noticeably different to normal phase GC × GC. In normal phase, the alkanes and iso-alkanes elute before the PAHs in the second dimension due to their low affinity for the polar column. In reversed phase, the alkanes elute after the PAHs in the second dimension and are shown as a band along the top of the contour plot (Fig. 4).
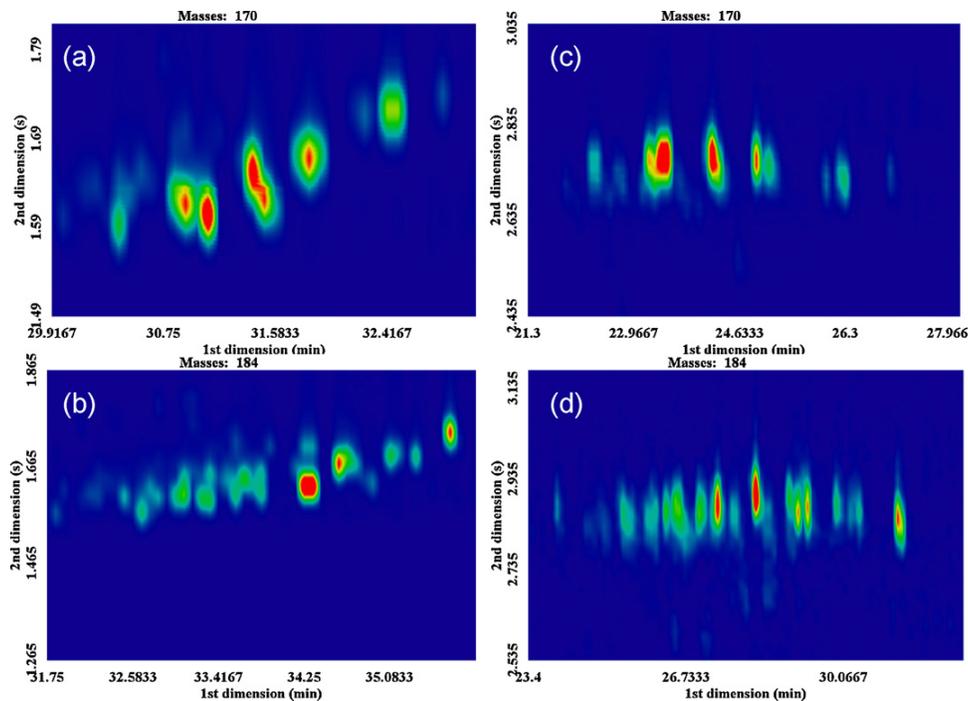
**Fig. 2.** GC × GC contour plots of C3 and C4 alkyl naphthalenes in DNAPL 1 using normal phase, (a) and (b) respectively, and reversed phase, (c) and (d) respectively.
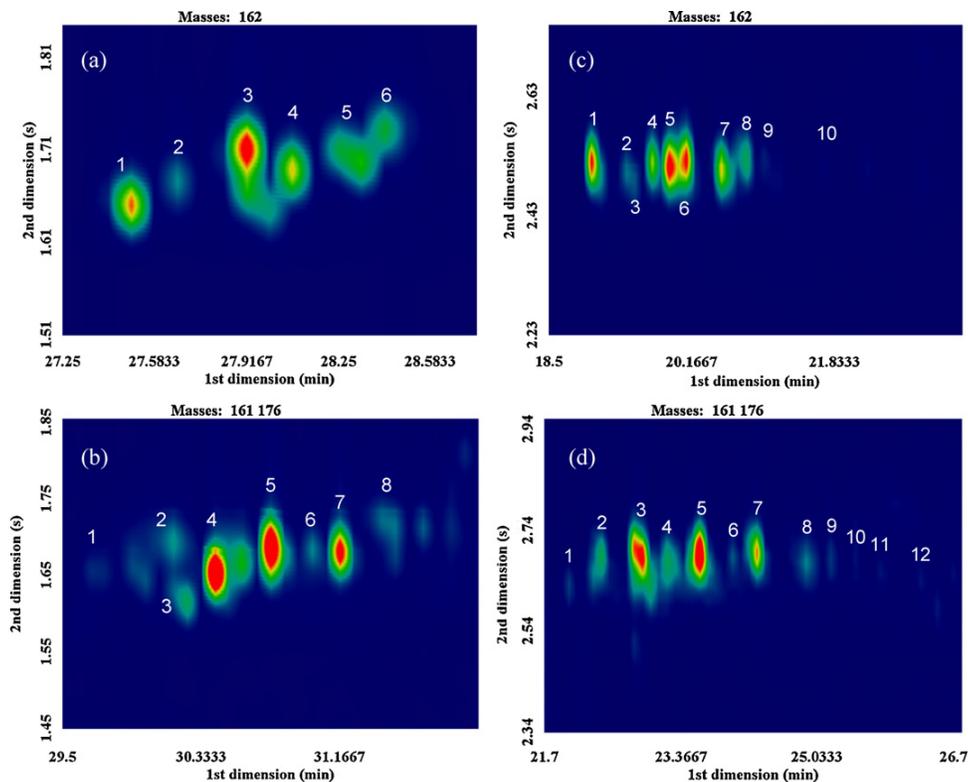


**Fig. 3.** GC × GC contour plots of C2 and C3 alkyl benzothiophenes in DNAPL 1 using normal phase, (a) and (b) respectively, and reversed phase, (c) and (d) respectively. Numbering indicates the peaks identified as alkyl benzothiophene isomers.

284
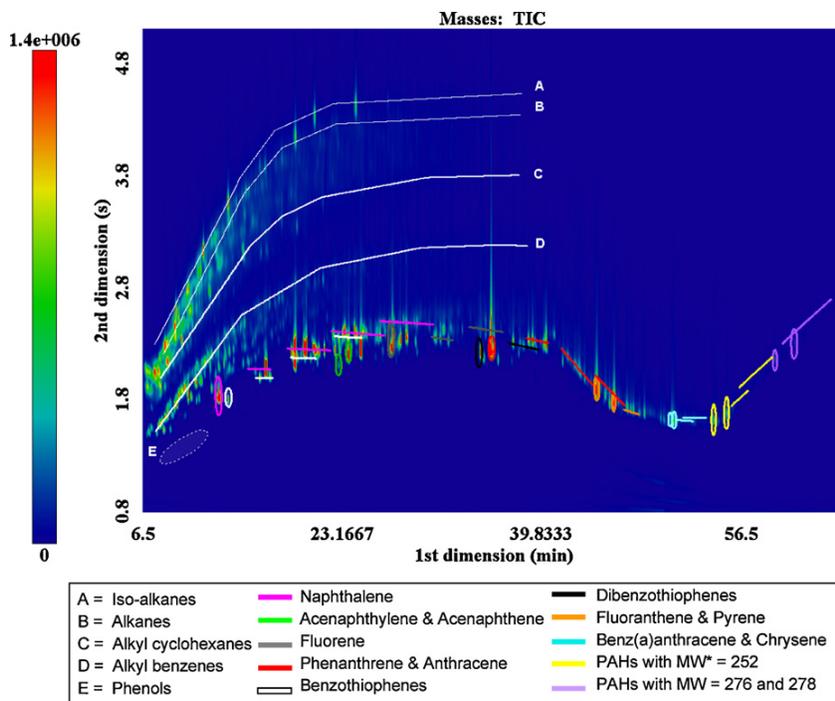
**Fig. 4.** GC × GC contour plot (in the total ion mode) of DNAPL 12. The key shows the identity of the compounds represented by each coloured circle. The solid lines of the same colour denote the alkylated derivatives of the corresponding compound [*MW = molecular weight].

### 3.3. PAH composition of DNAPLs

The 16 U.S. EPA priority PAHs were quantified using reversed phase GC × GC TOFMS. The PAH concentrations were corrected based on the percentage recovery values for the nearest eluting surrogate. The repeatability of the technique was evaluated by performing six identical injections of DNAPL sample 8. The average relative standard deviation (RSD) of the quantification was 3.0%, ranging from 0.3 to 4.8%. The low RSD values reflect the high separating power of GC × GC TOFMS.

Quantification was performed for the 16 EPA PAHs in all twelve DNAPL samples (Table 2). The majority of samples have very similar PAH fingerprints, with the same compounds being found in highest concentrations. Naphthalene was the most prevalent parent PAH in all DNAPLs except samples 11 and 12, where phenanthrene was found in highest concentrations. It is possible that this distinction is merely due to more advanced weathering in these samples (which will be discussed in more detail later).

Several PAH ratios were investigated as a simple method of comparing the DNAPL samples. The ratios used were Ant/(Ant + Phe), Flt/(Flt + Pyr) and BaA/(BaA + Chr), where Ant = anthracene, Phe = phenanthrene, Flt = fluoranthene, Pyr = pyrene, BaA = benz[a]anthracene and Chr = chrysene; the results are summarised in Table 3. These ratios have been used previously as a measure of pyrogenic/petrogenic character [35,36]. For pyrogenic samples, such as coal tar DNAPLs, Ant/(Ant + Phe), Flt/(Flt + Pyr) and BaA/(BaA + Chr) should give values greater than 0.10, 0.50 and 0.35 respectively [35,36]. However, for Flt/(Flt + Pyr) DNAPLs 1–5 produce values less than 0.50. Yunker et al. [36] state that values of 0.4–0.5 generally indicate combustion of liquid fossil fuel while values greater than 0.5 are indicative of solid fossil fuel combustion (e.g. coal). The unusual values for DNAPLs 1–5 could be explained by the presence of a gas reforming plant on site A, where petroleum fractions (instead of coal) were used to produce gas.

### 3.4. Chemical fingerprinting of DNAPLs

DNAPLs from different FMGP sites may differ widely in composition due to various factors involved in the manufacturing process. For example, different shapes of retort stand used to hold the coal during the carbonisation process will produce different by-products in the DNAPLs [28]. Similarly, low-temperature processes will produce DNAPLs containing a greater range of volatile components than high-temperature (>1000 °C) processes, as the higher temperatures tend to further degrade the volatile products [28].

However, PAH composition alone is not capable of differentiating between all seven FMGP sites. For the final part of this study, chemical fingerprints of the DNAPLs were produced by collating the peak data for a range of compounds, including aliphatics, alkylated PAHs and heterocyclic PAHs, using principal component analysis (PCA) for effective source differentiation.

Heterocyclic rings were the first group of chemicals to be explored, as compounds such as dibenzothiophenes are known to be resistant to environmental degradation processes and are used frequently for source identification of oil spills. GC × GC contour plots of the twelve DNAPLs allowed simple comparison of their chemical composition by visual inspection, allowing the major differences within the samples to be detected and investigated further using diagnostic ratio plots. A number of ratios were investigated using the peak areas (normalised to the internal standard) of carbazole (CBZ), dibenzofuran (DBF) and dibenzothiophene (DBT). The ratios CBZ/DBF and CBZ/DBT are represented as a cross plot in Fig. 5. The plot shows that the ratios can effectively separate the major types of manufacturing process, but are not capable of discerning between smaller differences, such as retort shape. The low CBZ/DBT values of sites containing a reforming gas plant may be due to the presence of petroleum fractions which would most likely contain high levels of dibenzothiophene with respect to carbazole.

Principal component analysis (PCA) was then performed to compare the chemical compositions of the twelve DNAPLs in an attempt

**Table 2**
PAH composition of DNAPLs (units are in mg/kg).

| Compounds | DNAPL | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
| Naphthalene | 16,797 | 9867 | 4617 | 5788 | 1068 | 481 | 47,171 | 85,432 | 31,763 | 81,931 | 36,169 | 4222 |
| Acenaphthylene | 4206 | 3891 | 1605 | 1857 | 503 | 252 | 8309 | 16,567 | 5044 | 7131 | 2333 | 924 |
| Acenaphthene | 883 | 407 | 244 | 252 | 93 | 57 | 1365 | 1485 | 1455 | 1538 | 22,458 | 6678 |
| Fluorene | 1917 | 1334 | 774 | 832 | 267 | 122 | 5612 | 6970 | 3644 | 5251 | 15,764 | 6877 |
| Phenanthrene | 3341 | 2730 | 1743 | 1930 | 567 | 271 | 21,304 | 29,445 | 10,892 | 20,107 | 36,896 | 11,361 |
| Anthracene | 3511 | 1308 | 945 | 984 | 437 | 235 | 11,298 | 7194 | 7180 | 8982 | 15,630 | 7774 |
| Fluoranthene | 1773 | 860 | 962 | 975 | 269 | 223 | 16,589 | 13,630 | 9220 | 14,988 | 10,518 | 8157 |
| Pyrene | 2079 | 1414 | 1074 | 1076 | 292 | 216 | 14,526 | 11,821 | 7165 | 12,844 | 7926 | 6141 |
| Benz[a]anthracene | 688 | 441 | 372 | 416 | 103 | 84 | 6763 | 5184 | 3265 | 5538 | 2228 | 4904 |
| Chrysene | 710 | 388 | 326 | 348 | 59 | 46 | 6674 | 4759 | 3706 | 5307 | 2447 | 4426 |
| Benzo[b]fluoranthene | 45 | 79 | 77 | 117 | 11 | 16 | 3187 | 1957 | 1656 | 2990 | 663 | 2472 |
| Benzo[k]fluoranthene | 306 | 174 | 170 | 226 | 45 | 32 | 4855 | 6408 | 2543 | 4362 | 828 | 2894 |
| Benzo[a]pyrene | 3282 | 1046 | 899 | 917 | 363 | 267 | 15,367 | 18,823 | 11,506 | 15,016 | 3137 | 3371 |
| Indeno[1,2,3-cd]pyrene | 42 | 83 | 43 | 79 | 13 | 13 | 2303 | 2133 | 1010 | 2165 | 306 | 1395 |
| Dibenz[a,h]anthracene | 196 | 48 | 23 | 29 | 15 | 14 | 606 | 913 | 490 | 600 | 62 | 868 |
| Benzo[g,h,i]perylene | 272 | 62 | 55 | 78 | 32 | 26 | 2263 | 2183 | 1271 | 2168 | 371 | 1332 |

**Table 3**

| DNAPL | Ant/(Ant + Phe) | Flt/(Flt + Pyr) | BaA/(BaA + Chr) |
|---|---|---|---|
| 1 | 0.51 | 0.46 | 0.49 |
| 2 | 0.32 | 0.38 | 0.53 |
| 3 | 0.35 | 0.47 | 0.53 |
| 4 | 0.34 | 0.48 | 0.54 |
| 5 | 0.44 | 0.48 | 0.64 |
| 6 | 0.46 | 0.51 | 0.65 |
| 7 | 0.35 | 0.53 | 0.50 |
| 8 | 0.20 | 0.54 | 0.52 |
| 9 | 0.40 | 0.56 | 0.47 |
| 10 | 0.31 | 0.54 | 0.51 |
| 11 | 0.30 | 0.57 | 0.48 |
| 12 | 0.41 | 0.57 | 0.53 |

to fully differentiate between the manufacturing processes used at the seven FMGP sites. In total, the data for 140 peaks was entered into the software, including PAHs and their alkyl homologues, alkanes, alkyl benzenes and a range of heterocyclic PAH compounds, resulting in the PCA score plot shown in Fig. 6a. A full list of the compounds used to prepare each score plot has been included in the supplementary information (Table S1). The inclusion of peak areas for individual alkylated PAHs and heterocyclic PAHs provides 77 more data points per sample than if groupings by alkylation level were used, as in conventional chemical fingerprinting methods.

The first two principal components in Fig. 6a describe 78.5% of the total variation in the data set. The score plot illustrates that it



**Fig. 6.** Principal component analysis (PCA) score plots comparing peak areas of (a) 140 different components* and (b) compounds susceptible to weathering, found in twelve DNAPL samples (labelled 1–12) from seven different FMGP sites. [*a full list of compounds used to prepare each plot can be found in Table S1 of the supplementary information].

is possible to distinguish between the FMGP sites based on DNAPL composition. DNAPLs 1–6 originate from the same FMGP site (site A) and are shown to be very similar in nature as they form a cluster in the score plot. This indicates that the peak data used to prepare the score plot provides a good source fingerprint for the DNAPL samples. Samples 1–6 are clustered in a separate quadrant to the other coal retort DNAPLs. It is hypothesised that the variety of processes used at site A has resulted in a complex mix of coal retort tar and reforming gas plant contamination.



**Fig. 5.** Cross plot of heterocyclic ratios using carbazole (CBZ), dibenzofuran (DBF) and dibenzothiophene (DBT) for source apportionment.

Horizontal and vertical retorts create different by-products as their different shapes cause the gases evolved during the carbonisation of coal to be kept in contact with the hot walls of the retort for different lengths of time. Previous literature states that DNAPLs produced by a horizontal retort will be rich in phenol and naphthalene [28], most likely due to the increased contact time with the retort walls allowing the higher molecular weight PAHs to be degraded. Sample 8 appears in a different quartile of the score plot to the other horizontal retort DNAPLs (samples 7 and 10). This sample was obtained from inside a tar tank, thus it is likely that the differences in chemical composition may be due to a lower extent of weathering than in the other samples.

As expected, DNAPL samples 11 and 12 showed a high degree of difference from the other samples as they originated from a wood preservative site and water gas site respectively, whereas all other samples were obtained from horizontal or vertical retort coal gasworks. The DNAPL found at a wood preservative site is most likely to be from creosote oil, a distilled fraction of coal tar DNAPLs which was used to treat wood [28]; hence samples from such sites will likely exhibit a smaller range of compounds than those obtained from water gas and coal retort sites. Alkane peak areas were included in this plot as previous literature [28] states that they are more prevalent in DNAPLs from water gas sites, allowing sample 12 to be easily distinguished from the other DNAPLs.

### 3.5. Weathering of DNAPLs

The use of a wide variety of compounds in the initial PCA score plot (Fig. 6a) provided a chemical fingerprint able to distinguish between different DNAPL sources. A further PCA score plot was prepared in an attempt to differentiate between the samples based on the degree of weathering present (Fig. 6b). The 'weathering plot' incorporates a number of weathering ratios calculated for each DNAPL and the peak areas of low molecular weight compounds which are most susceptible to weathering. A weathering ratio for alkanes, using the peak areas of the straight chain alkanes (or n-alkanes) divided by the peak areas of the branched alkanes, was included to ensure the large differences in alkane concentrations caused by different manufacturing processes was not an issue. The values for a PAH weathering ratio, calculated based on the total C2 and C3 alkyl naphthalenes (C2N and C3N respectively) compared to naphthalene (C0N) itself, C0N/(C2N + C3N), was also included due to the tendency for alkyl homologues to be more prevalent in severely weathered samples.[7] Furthermore, an equivalent ratio for the benzothiophenes (one of the most prevalent heterocyclic families found in the DNAPLs) was also included. The calculated ratios for each sample are given in the supplementary information (Table S4).

The first two principal components in the weathering plot describe 82.5% of the total variation. Samples 1–6 show a more pronounced difference in this plot compared to the initial score plot in Fig. 6a. This demonstrates that the there is a degree of difference in their chemical fingerprints which may be attributed to different weathering processes occurring across the site. Samples 1 and 2 were obtained from boreholes in a similar area at site A, whereas samples 3–5 were all obtained from within a tar tank near the site boundary and sample 6 was obtained from a borehole next to the same tar tank. This is illustrated in the weathering plot, as samples 1 and 2 have separated from the cluster of other samples obtained from site A.

By inspection of the peak areas and ratios used to prepare the PCA plot in Fig. 6b, a trend in the degree of weathering can be approximated; samples in the top-right quartile appear to show less severe weathering, while those towards the bottom-left quartile indicate the most severe cases of weathering. The degree of weathering in samples 11 and 12 was evaluated by inspec-

tion of the concentrations of alkyl PAHs relative to the parent PAH. The concentration of PAHs and their homologues in the majority of samples display the expected pyrogenic PAH pattern, C0 > C1 > C2 > C3, signifying that these samples have not undergone significant weathering [37]. However, in samples 11 and 12 (and to a slightly lesser extent in samples 5 and 6) the concentrations are more similar to the accepted weathering pattern of C0 < C1 < C2 < C3, indicating that these samples are more severely weathered. An illustration of the weathering patterns of naphthalenes and benzothiophenes in each of the DNAPL samples can be found in the supplementary information (Figs. S5 and S6).

PCA plots have been shown to be capable of not only distinguishing between DNAPLs from different types of FMGP sites, but the degree of weathering can also be estimated by exclusion of the more stable compounds generally used as source indicators. This analytical process could prove very useful in distinguishing the differences between DNAPLs caused by differences in the manufacturing processes employed at FMGP sites, as well as in distinguishing differences in chemical fingerprint across a single FMGP site. As previously mentioned, many FMGP sites have now been split into various land holdings and PCA plots of GC × GC data from across the entire site may help to indicate the presence of multiple sources of contamination, thus determining the persons liable for remediation costs. The reasons for variation of chemical fingerprint across a site, e.g. environmental weathering or multiple contamination sources, can be confirmed using ancillary methods, such as CSIA. The combination of reversed phase GC × GC with PCA outlined in this study allows large amounts of chemical information to be generated for each sample but collated in a manageable format. It is for this reason that we deem this method of chemical fingerprinting as "ultra resolution".

## 4. Conclusions

This study details the first attempt at development of a standard approach to chemical fingerprinting of coal tar DNAPLs. Conventional tiered approaches to chemical fingerprinting involve tedious sample preparation and cleanup steps, multiple analytical instruments and complicated data processing. The use of reversed phase GC × GC TOFMS provides an accurate and precise method of chemical fingerprinting for complex samples, such as coal tar, by analysis of a single, non-specific sample extract using a single analytical instrument. The application of principal component analysis to sections of the GC × GC dataset has been shown to simplify the comparison of highly complex samples. PCA score plots can be used to compare the chemical fingerprints of a number of samples at once, allowing site-specific differences to be easily identified. The method described could prove particularly useful for source identification and monitoring of natural attenuation during environmental forensic investigations at former gasworks and at a multitude of other contaminated sites.

## Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.chroma.2011.05.045.

## References

[1] F. D'Affonseca, P. Blum, M. Finkel, R. Melzer, P. Grathwohl, J. Contam. Hydrol. 102 (2008) 120.
[2] P.S. Birak, C.T. Miller, J. Contam. Hydrol. 105 (2009) 81.
[3] D.G. Brown, L. Gupta, T.H. Kim, H.K. Moo-Young, A.J. Coleman, Chemosphere 65 (2006) 1562.
[4] Z. Wang, S.A. Stout, in: Z. Wang, S.A. Stout (Eds.), Oil Spill Environmental Forensics: Fingerprinting and Source Identification, Elsevier, London, 2007, p. 1.
[5] Department of Environment, Food and Rural Affairs (DEFRA), Directive 2004/35/CE of the European Parliament and Council of the European Union on environmental liability with regard to the prevention and remedying of environmental damage, 2009. http://www.defra.gov.uk/environment/policy/liability/.
[6] P.S. Daling, L.G. Faksness, A.B. Hansen, S.A. Stout, Environ. Forensic 3 (2002) 263.
[7] G.S. Douglas, S. Emsbo-Mattingly, S.A. Stout, A.D. Uhler, K.J. McCarthy, in: B.L. Murphy, R.D. Morrison (Eds.), Introduction to Environmental Forensics, Elsevier, London, 2007, p. 312.
[8] Z.D. Wang, M. Fingas, M. Landriault, L. Sigouin, Y. Feng, J. Mullin, J. Chromatogr. A 755 (1997) 251.
[9] R.B. Gaines, G.S. Frysinger, C.M. Reddy, R.K. Nelson, in: Z. Wang, S.A. Stout (Eds.), Oil Spill Environmental Forensics: Fingerprinting and Source Identification, Elsevier, London, 2007, p. 169.
[10] P.D. Boehm, in: R.D. Morrison, B.L. Murphy (Eds.), Environmental Forensics: Contaminant Specific Guide, Elsevier, London, 2006, p. 314.
[11] F. Haeseler, D. Blanchet, V. Druelle, P. Werner, J.P. Vandecasteele, Environ. Sci. Technol. 33 (1999) 825.
[12] D. Saber, D. Mauro, T. Sirivedhin, Environ. Forensic 7 (2006) 65.
[13] C. Eberhardt, P. Grathwohl, J. Contam. Hydrol. 59 (2002) 45.
[14] A.W. Hatheway, Eng. Geol. 64 (2002) 317.
[15] A.O. Thomas, J.N. Lester, Sci. Total Environ. 152 (1994) 239.
[16] O. Panić, T. Górecki, Anal. Bioanal. Chem. 386 (2006) 1013.
[17] M. Adahchour, J. Beens, R.J.J. Vreuls, U.A. Brinkman, TrAC, Trends Anal. Chem. 25 (2006) 438.
[18] M. Adahchour, J. Beens, R.J.J. Vreuls, U.A. Brinkman, TrAC, Trends Anal. Chem. 25 (2006) 540.
[19] J. Dallüge, J. Beens, U.A. Brinkman, J. Chromatogr. A 1000 (2003) 69.
[20] P. Marriott, R. Shellie, TrAC, Trends Anal. Chem. 21 (2002) 573.
[21] J.B. Phillips, J. Beens, J. Chromatogr. A 856 (1999) 331.
[22] W. Bertsch, J. High Resolut. Chromatogr. 22 (1999) 647.
[23] R. Van Der Westhuizen, A. Crouch, P. Sandra, J. Sep. Sci. 31 (2008) 3423.
[24] G.S. Frysinger, R.B. Gaines, C.M. Reddy, Environ. Forensic 3 (2002) 27.
[25] E. Skoczyńska, P. Korytár, J. De Boer, Environ. Sci. Technol. 42 (2008) 6611.
[26] C. Muhlen, C.A. Zini, E.B. Caramao, P.J. Marriot, J. Chromatogr. A 1105 (2006) 39.
[27] A. Findlay, The Treasures of Coal Tar, Turnbull and Spears, UK, 1917.
[28] T.H. Butler, in: S. Young (Ed.), Distillation Principles and Processes, Macmillan and Co, UK, 1922, p. 359.
[29] U.S. Environmental Protection Agency, Priority Pollutants http://water.epa.gov/scitech/swguidance/methods/pollutants.cfm [accessed 22.10.10].
[30] J.H. Christensen, G. Tomasi, J. Chromatogr. A 1169 (2007) 1.
[31] Z. Wang, M. Fingas, Mar. Pollut. Bull. 47 (2003) 423.
[32] A. Assal, 2009. 'Environmental forensics: compound specific isotope analysis of PAHs. Study of a former tar plant', PhD thesis, Queen's University, Belfast.
[33] U.S. Environmental Protection Agency, Test Methods for Evaluating Solid Wastes, SW-846 Method 8000B. http://www.epa.gov/waste/hazard/testmethods/sw846/pdfs/8000b.pdf.
[34] H. Van De Weghe, G. Vanermen, J. Gemoets, R. Lookman, D. Bertels, J. Chromatogr. A 1137 (2006) 91.
[35] H. Budzinski, I. Jones, J. Bellocq, C. Pierard, P. Garrigues, Mar. Chem. 58 (1997) 85.
[36] M.B. Yunker, R.W. Macdonald, R. Vingarzan, R.H. Mitchell, D. Goyette, S. Sylvestre, Org. Geochem. 33 (2002) 489.
[37] Z. Wang, S.A. Stout, in: R.E. Hester, R.M. Harrison (Eds.), Issues in Environmental Science and Technology, Royal Society of Chemistry London, 2008, p. 54.